

МОДЕЛИРОВАНИЕ СЛОЖНЫХ ВЕРОЯТНОСТНЫХ СИСТЕМ

Министерство образования и науки Российской Федерации
Уральский федеральный университет
имени первого Президента России Б.Н. Ельцина

МОДЕЛИРОВАНИЕ СЛОЖНЫХ ВЕРОЯТНОСТНЫХ СИСТЕМ

Учебное пособие

*Рекомендовано Региональным отделением УрФО учебно-методического
объединения вузов Российской Федерации по образованию в области
радиотехники, электроники, биомедицинской техники и автоматизации в
качестве учебного пособия для студентов высших учебных заведений,
обучающихся по направлению подготовки
220400 – Управление в технических системах в УрФО*

Екатеринбург
УрФУ
2011

УДК 681.5:001.891.57(075.8)
ББК 32.817в6я73
М74

Авторы: В.Г. Лисиенко, О.Г. Трофимова, С.П. Трофимов, Н.Г. Дружинина,
П.А. Дюгай.

Рецензенты:

кафедра информационных технологий Института урбанистики Уральской государственной архитектурно-художественной академии (канд. физ.-мат. наук А.И. Кривоногов); начальник группы отд. 315 НПО «Автоматика», ст. науч. сотр., канд. техн. наук Г.П. Лосев

Научный редактор доц., канд. техн. наук В.А. Морозова

М74 Моделирование сложных вероятностных систем : учеб. пособие /
В. Г. Лисиенко, О. Г. Трофимова, С. П. Трофимов, Н. Г. Дружинина,
П.А. Дюгай. Екатеринбург: УРФУ, 2011. 200 с.

ISBN 978-5-321-02060-9

В данном издании рассмотрены различные виды статистического анализа и обработки результатов стохастического моделирования. В практической части пособия приведены характерные примеры вероятностных моделей: инженерный вероятностный калькулятор, игровой автомат, приложение «Узоры», вероятностная логистическая модель информационно-коммуникационной транспортной системы, рассмотрены авторские олимпиадные задачи по теории вероятностей с решением. Учебное пособие предназначено для студентов, обучающихся по направлениям подготовки инженеров, бакалавров и магистров 220400 – Управление в технических системах и 230100 – Информатика и вычислительная техника, а также для специалистов, занимающихся вопросами управления и моделирования сложных вероятностных систем.

Библиогр.: 42 назв. Табл. 24. Рис. 86.

УДК 681.5:001.891.57(075.8)
ББК 32.817в6я73

ISBN 978-5-321-02060-9

© Уральский федеральный университет, 2011
© Лисиенко В.Г., Трофимова О.Г.,
Трофимов С.П., Дружинина Н.Г.,
Дюгай П.А., 2011

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ	5
ЧАСТЬ 1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ МОДЕЛИРОВАНИЯ МАТЕМАТИЧЕСКИХ СИСТЕМ	
1. МЕТОДОЛОГИЯ МОДЕЛИРОВАНИЯ СИСТЕМ.....	7
1.1. МОДЕЛЬ И ВИДЫ МОДЕЛИРОВАНИЯ	7
1.2. РАЗВИТИЕ ОПРЕДЕЛЕНИЯ СИСТЕМЫ	10
1.3. ПОНЯТИЯ, ХАРАКТЕРИЗУЮЩИЕ ФУНКЦИОНИРОВАНИЕ СИСТЕМЫ.....	13
2. КЛАССИФИКАЦИЯ МОДЕЛЕЙ	19
2.1. КЛАССИФИКАЦИЯ МОДЕЛЕЙ И МЕТОДОВ ФОРМАЛИЗОВАННОГО ПРЕДСТАВЛЕНИЯ СИСТЕМ.....	19
2.2. КЛАССИФИКАЦИЯ ВЕРОЯТНОСТНЫХ МОДЕЛЕЙ	36
3. ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ.....	43
3.1. ОСНОВНЫЕ ЭТАПЫ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ.....	44
3.2. ИДЕНТИФИКАЦИЯ И ВЕРИФИКАЦИЯ ИМИТАЦИОННОЙ МОДЕЛИ.....	50
3.3. ПЛАНИРОВАНИЕ ИМИТАЦИОННЫХ ЭКСПЕРИМЕНТОВ.....	53
4. КОРРЕЛЯЦИОННЫЙ АНАЛИЗ	56
4.1. КОРРЕЛЯЦИЯ И ПРИЧИННАЯ ЗАВИСИМОСТЬ	56
4.2. КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ	57
4.3. ИНТЕРПРЕТАЦИЯ КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ	62
5. РЕГРЕССИОННЫЙ АНАЛИЗ	64
5.1. МЕТОД НАИМЕНЬШИХ КВАДРАТОВ	64
5.2. ВЫБОР ФОРМЫ ФУНКЦИОНАЛЬНОЙ ЗАВИСИМОСТИ.....	68
5.3. ЛИНЕЙНЫЕ РЕГРЕССИОННЫЕ УРАВНЕНИЯ	71
6. СТАТИСТИЧЕСКОЕ ОЦЕНИВАНИЕ И ПРОВЕРКА ГИПОТЕЗ	75
6.1. ТЕОРЕТИЧЕСКИЕ РАСПРЕДЕЛЕНИЯ, ИСПОЛЬЗУЕМЫЕ В СТАТИСТИЧЕСКИХ ВЫВОДАХ	75
6.1.1. Распределение хи-квадрат	77
6.1.2. F-распределение	79
6.1.3. Распределение Стьюдента	79
6.1.4. Биномиальное распределение	80
6.1.5. Распределение Пуассона	81
6.1.6. Распределение Эрланга.....	84
6.2. ВЫБОРОЧНЫЕ РАСПРЕДЕЛЕНИЯ	85
6.3. ИДЕНТИФИКАЦИЯ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ.....	88
6.3.1. Использование метода наименьших квадратов	88
6.3.2. Использование числовых характеристик выборки	89
6.4. ГЕНЕРАЦИЯ СЛУЧАЙНЫХ ВЕЛИЧИН	92
6.4.1. Метод «обратной» функции	92
6.4.2. Генерация двумерных случайных величин	93
6.5. ИНТЕРВАЛЬНОЕ ОЦЕНИВАНИЕ	95

6.6. ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ	99
7. ПРИМЕНЕНИЕ КРИТЕРИЕВ СОГЛАСИЯ.....	104
7.1. КРИТЕРИИ СОГЛАСИЯ ДЛЯ СРЕДНИХ.....	105
7.2. КРИТЕРИИ СОГЛАСИЯ ДЛЯ ДИСПЕРСИЙ.....	112
7.3. КРИТЕРИИ СОГЛАСИЯ ДЛЯ КОРРЕЛЯЦИОННЫХ ПОКАЗАТЕЛЕЙ.....	115
7.4. КРИТЕРИИ ДЛЯ ПРОВЕРКИ СЛУЧАЙНОСТИ И ОЦЕНКИ РЕЗКО ВЫДЕЛЯЮЩИХСЯ НАБЛЮДЕНИЙ.....	117
8. ОСНОВЫ ДИСПЕРСИОННОГО АНАЛИЗА	120
8.1. ДВЕ МОДЕЛИ ДИСПЕРСИОННОГО АНАЛИЗА	121
8.2. РАЗБИЕНИЕ СУММЫ КВАДРАТОВ.....	123
8.3. ОДНОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ С ПОСТОЯННЫМИ ЭФФЕКТАМИ.....	124
8.4. МНОЖЕСТВЕННОЕ СРАВНЕНИЕ СРЕДНИХ	127
КОНТРОЛЬНЫЕ ВОПРОСЫ	128
ЧАСТЬ 2. ПРИМЕРЫ ВЕРОЯТНОСТНЫХ МОДЕЛЕЙ	
9. ВЕРОЯТНОСТНЫЙ КАЛЬКУЛЯТОР ДЛЯ ИНЖЕНЕРНЫХ РАСЧЕТОВ	130
10. АНАЛИЗ ОБМЕННЫХ КУРСОВ ВАЛЮТ В БАНКАХ ГОРОДА.....	133
11. МОДЕЛИРОВАНИЕ ИГРОВОГО АВТОМАТА	135
12. МНОГОПАРАМЕТРИЧЕСКАЯ СТОХАСТИЧЕСКАЯ МОДЕЛЬ «УЗОРЫ»	140
12.1. ИЗОБРАЖЕНИЯ С КОНЕЧНЫМ НАБОРОМ ДАННЫХ	140
12.2. УЗОРЫ.....	141
12.2.1. Геометрические фракталы.....	142
12.2.2. L-системы.....	143
12.2.3. Алгебраические фракталы.....	145
12.2.4. Стохастические фракталы	146
12.3. ЛИНЕЙЧАТЫЕ ИЗОБРАЖЕНИЯ	146
12.3.1. Обобщенные многоугольники	146
12.3.2. Аналитический подход к построению линейчатых узоров	148
13. ОЛИМПИАДНЫЕ ЗАДАЧИ И ИХ РЕШЕНИЯ.....	153
14. ИНФОРМАЦИОННО-КОММУНИКАЦИОННАЯ ТРАНСПОРТНАЯ СИСТЕМА	160
14.1. ЛОГИСТИЧЕСКАЯ МОДЕЛЬ ТРАНСПОРТНОГО ПРЕДПРИЯТИЯ НА ОСНОВЕ ИНФОРМАЦИОННО- КОММУНИКАЦИОННОЙ СИСТЕМЫ	160
14.2. УПРАВЛЕНИЕ И АНАЛИЗ РАБОТЫ ТРАНСПОРТНОЙ СЕТИ С ПОМОЩЬЮ СПУТНИКОВОЙ НАВИГАЦИИ	184
14.3. ДИСПЕРСИОННЫЙ АНАЛИЗ ТЕХНИКО-ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ ЛОГИСТИЧЕСКОЙ МОДЕЛИ ТРАНСПОРТНОЙ СЕТИ	189
ЗАКЛЮЧЕНИЕ	193
ОТВЕТЫ НА КОНТРОЛЬНЫЕ ВОПРОСЫ	194
БИБЛИОГРАФИЧЕСКИЙ СПИСОК	197

ПРЕДИСЛОВИЕ

Данное учебное пособие посвящено математическому моделированию вероятностных систем и состоит из двух частей: теоретической и практической.

Современные образовательные стандарты предлагают модульный подход при изучении дисциплин. В связи с этим учебное пособие ставит своей целью развитие способности у студентов применять знания теории вероятностей, математической статистики и случайных процессов и других дисциплин, в частности основ математического моделирования систем, управления вычислительным экспериментом с моделью, обработки результатов исследований и др. Это позволит сформировать профессиональные компетенции, такие как способность выполнять эксперименты на действующих объектах по заданным методикам и обрабатывать результаты с применением современных информационных технологий и технических средств, а также способность проводить вычислительные эксперименты с использованием стандартных программных средств с целью получения математических моделей процессов и объектов автоматизации и управления.

В теоретической части пособия приводятся основные понятия, связанные с моделированием систем (модель, система, цель, структура, состояние и т.п.), дается классификация моделей и методов формализованного представления систем, в том числе и вероятностных моделей, а также рассматриваются этапы имитационного моделирования. В частности, подробно рассмотрены различные виды статистического анализа и обработки результатов стохастического моделирования. Основой любого анализа экспериментальных данных является решение задач идентификации параметров распределений, статистического оценивания и проверка гипотез с применением критериев согласия, поэтому в теоретической части учебного пособия этим задачам уделяется определенное внимание.

В практической части пособия приведены характерные примеры вероятностных моделей.

В данном издании описывается библиотека иерархии классов, позволяющих работать со случайными величинами как с объектами. Для этих объектов переопределены арифметические операции и основные математические функции. При изучении дисциплин, связанных с вероятностными расчетами, предлагается использовать уникальный инженерный вероятностный калькулятор.

Практический интерес для популяризации знаний о вероятностных подходах представляет анализ обменных курсов валют в банках города. Доступная для общественности информация о курсах покупки/продажи валюты дополняется аналитическими графиками, которые наглядно показывают текущие тенденции на валютном рынке.

Актуальной задачей является построение вероятностной модели технической системы, основанной на многократном использовании некоторой случайной величины. В качестве примера такой модели рассматривается игровой автомат. Предложенные уравнения позволяют управлять основными характеристиками модели автомата – доходностью и риском.

Математический аппарат приложения «Узоры» позволяет моделировать и распознавать регулярные линейчатые узоры на плоскости, проводить идентификацию параметров узоров. На базе этого приложения можно проводить многоплановые эксперименты и статическое оценивание результатов исследования.

Кроме того, рассмотрены авторские олимпиадные задачи по теории вероятностей с решением. Эти задачи демонстрируют полезные приемы, которые могут использоваться в сложных вероятностных моделях. В частности, рассматриваются вопросы генерации случайных величин, операции и функции случайных аргументов. Ряд задач посвящен стохастическим финансовым операциям, использующимся в повседневной жизни. Неформальные постановки задач вырабатывают у студентов общекультурные компетенции в овладении культурой мышления, в способности к обобществлению, анализу, восприятию информации, постановке цели и выбору путей ее достижения.

В заключительной главе представлена вероятностная логистическая модель оригинальной информационно-коммуникационной транспортной системы как пример сложной вероятностной модели, позволяющей провести многокритериальный статистический анализ, обработку и интерпретацию результатов исследований.

Таким образом, практическая часть пособия позволяет студентам приобрести и закрепить навыки теоретического освоения материала первой части пособия, в частности, формализовать модель, провести запланированный эксперимент, оценить, проанализировать, обработать и идентифицировать результаты исследований согласно заданным целям и критериям моделирования. Нестандартные формулировки задач помогут студентам расширить круг предметных областей, в которых возможно применение вероятностных моделей.

Учебное пособие может быть использовано при изучении целого комплекса дисциплин: «Теория вероятности и математическая статистика», «Моделирование систем управления», «Теория и методы управления экспериментом», «Идентификация и диагностика систем», «Системы массового обслуживания», «Моделирование информационных систем и сетей», «Информационное обеспечение систем управления», «Автоматизированные информационно-управляющие системы», «Технологии управления и методы анализа результатов экспериментальных исследований».

Авторы будут благодарны всем читателям, уделившим внимание данному изданию, а также выразившим свои возможные пожелания и замечания.

Авторы приносят благодарность руководству и издательскому отделу УрФУ им. первого Президента России Б.Н. Ельцина, а также руководству Института радиоэлектроники и информационных технологий – РТФ и коллективам кафедр автоматики и управления в технических системах и автоматики и информационных технологий за проявленное внимание и поддержку.

ЧАСТЬ 1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ МОДЕЛИРОВАНИЯ МАТЕМАТИЧЕСКИХ СИСТЕМ

1. МЕТОДОЛОГИЯ МОДЕЛИРОВАНИЯ СИСТЕМ

1.1. МОДЕЛЬ И ВИДЫ МОДЕЛИРОВАНИЯ

Понятие «модель» возникло в процессе опытного изучения мира. Слово «модель» произошло от латинского слова «modelium», означает: мера, образ, способ и т.д. Его первоначальное значение было связано со строительным искусством, и почти во всех европейских языках оно употреблялось для обозначения образа, или прообраза, или вещи, сходной в каком-то отношении с другой вещью [1].

Одна из первых попыток обобщенного определения модели была предпринята Г. Клаусом. Под моделью он понимает отображение фактов, вещей, отношений определенной области знаний в виде более простой и более наглядной материальной структуры. Главным требованием к модели он выдвинул «существенность сходства и несущественность различия с оригиналом» [2]. Кроме того, в моделировании сложных динамических систем можно достичь сходных результатов на основе сходства поведения, сходства структуры и материала, из которого они состоят.

Под моделью понимается такая мысленно представляемая или материально реализованная система, которая отображает и воспроизводит объект так, что ее изучение дает новую информацию об этом объекте [1].

Модель – это система, исследование которой служит средством для получения информации о другой системе [3].

Выделяют следующие обобщенные признаки модели [3]:

1. Модель не может существовать изолированно, потому что она всегда связана с оригиналом, т.е. той материальной или идеальной системой, которую она замещает в процессе познания.
2. Модель должна быть не только сходна с оригиналом, но и отлична от него, причем модель отражает те свойства и отношения оригинала, которые существенны для того, кто ее применяет.
3. Модель обязательно имеет целевое назначение. Таким образом, модель – это упрощенный образ оригинала, неразрывно с ним связанный, отражающий существенные свойства, связи и отношения оригинала; система, исследование которой служит инструментом, средством для получения новой и (или) подтверждения уже имеющейся информации о другой системе.

В энциклопедическом словаре «Информатика»¹ под моделью объекта понимается «другой объект (реальный, знаковый или воображаемый), отличный от исходного, который обладает существенными для целей

¹ Воройский Ф.С. Информатика. Энциклопедический словарь-справочник. М.: ФИЗМАТЛИТ, 2006. 768 с.

моделирования свойствами и в рамках этих целей полностью заменяет исходный объект».

Модель – это упрощенное описание сложной системы, в некоторых существенных частях являющееся изоморфным (т.е. подобным) оригинальной системе и отражающее существенные свойства системы. Такое подобие и способность воспроизводить эффекты оригинала позволяет использовать модель в качестве заместителя изучаемой системы.

Еще один вариант происхождения понятия «модель» от латинского «*modulus*», означающий «уменьшенный» вариант изначального способа. Этот термин является однокоренным с латинским словом «*modus*», которое означает «образ действий либо существования, метод, форма, манера, привычка, способ или стиль».

Таким образом, модель объекта обычно представляет собой миниатюрную версию или репрезентацию этого объекта. Действующая модель (например, машины) обозначает нечто, способное выполнять ту же работу, что и сама машина, но в меньшем объеме.

Понятие «модель» со временем стало обозначать также описание или аналогию, используемую для того, чтобы облегчить визуализацию чего-либо (например, атома), недоступного непосредственному наблюдению. Оно может использоваться также для обозначения формального описания некоего явления или положения вещей.

В [4] представлено 6 вариантов определений термина «модель»: 1) образец (эталон, стандарт) изделия для серийного производства, тип, марка изделия; 2) образец изделия, служащий для изготовления формы при отливке и воспроизведении в другом материале; 3) натурщик, позирующий художнику, изображаемые художником предметы («натура»); 4) устройство, воспроизводящее, имитирующее строение и действие какого-либо другого («моделируемого») приспособления в научных, производственных (при испытаниях) или спортивных целях; 5) схема, изображение или описание какого-либо явления или процесса в природе и обществе; 6) модель в языкознании – абстрактное понятие эталона какой-либо системы (фонетической, грамматической и т. п.), ее общих характеристик; общая схема описания языка или какой-либо его подсистемы.

Моделирование – это процесс и результат создания модели. Метод моделирования в современных условиях приобретает новые черты, значительно видоизменяясь и обобщаясь. В силу этого важным является формулирование его обобщенной характеристики. Моделирование – это метод опосредованного практического или теоретического оперирования объектом, при котором исследуется непосредственно не сам интересующий нас объект, а специально созданная вспомогательная искусственная или естественная система. При этом она должна находиться в определенном объективном соответствии с познаваемым объектом, способна замещать его на определенных этапах познания и давать при её исследовании в итоге информацию о самом моделируемом объекте. В этом состоит отличительная черта, присущая только процессу моделирования и отличающая его от других методов познания [5].

Моделирование можно рассматривать как главный способ познания окружающего мира. Определяя гносеологическую роль моделирования, отметим многообразие моделей в науке и технике. Моделируемый объект называется оригиналом, моделирующий – моделью.

Применительно к техническим системам *моделирование* представляет собой процесс замещения объекта исследования некоторой его моделью и проведение исследований на модели с целью получения необходимой информации об объекте или системе. *Модель* – это физический или абстрактный образ моделируемого объекта, удобный для проведения исследований и позволяющий адекватно отображать интересующие исследователя физические свойства и характеристики объекта. Удобство проведения исследований может определяться различными факторами: легкостью и доступностью получения информации, сокращением сроков и уменьшением материальных затрат на исследование [6].

Различают моделирование предметное и абстрактное. При *предметном моделировании* строят *физическую модель*, которая соответствующим образом отображает основные физические свойства и характеристики моделируемого объекта. При этом модель может иметь иную физическую природу в сравнении с моделируемым объектом (например, электронная модель гидравлической или механической системы). Если модель и объект одной и той же физической природы, то моделирование называют *физическим*.

Физическое моделирование широко применялось до недавнего времени при создании сложных технических объектов. Обычно изготавливался макетный или опытный образец технического объекта, проводились испытания, в процессе которых определялись его выходные параметры и характеристики, оценивались надежность функционирования и степень выполнения технических требований, предъявляемых к объекту. Если вариант технической разработки оказывался неудачным, все повторялось сначала, т. е. осуществлялось повторное проектирование, изготовление опытного образца, испытания и т. д.

Физическое моделирование сложных технических систем сопряжено с большими временными и материальными затратами.

Абстрактное моделирование связано с построением *абстрактной модели*. Такая модель представляет собой математические соотношения, графы, схемы, диаграммы и т. п. Наиболее мощным и универсальным методом абстрактного моделирования является математическое моделирование. Оно широко используется как в научных исследованиях, так и при проектировании.

Математическое моделирование позволяет посредством математических символов и зависимостей составить описание функционирования технического объекта в окружающей внешней среде, определить выходные параметры и характеристики, получить оценку показателей эффективности и качества, осуществить поиск оптимальной структуры и параметров объекта. Применение математического моделирования при проектировании в большинстве случаев позволяет отказаться от физического моделирования, значительно сократить объемы испытаний и доводочных работ, обеспечить создание технических

объектов с высокими показателями эффективности и качества. Одним из основных компонентов системы проектирования в этом случае становится математическая модель.

Математическая модель – это совокупность математических объектов и отношений между ними, адекватно отображающая физические свойства создаваемого технического объекта. В качестве математических объектов выступают числа, переменные, множества, векторы, матрицы и т. п. Процесс формирования математической модели и использования ее для анализа и синтеза называется *математическим моделированием*. В конструкторской практике под математическим моделированием обычно понимается процесс построения математической модели, а проведение исследований на модели в процессе проектирования называют *вычислительным (машинным или компьютерным) экспериментом*.

На различных этапах моделирования сложной технической системы используются различные математические модели. Вначале обычно строят простые модели, при дальнейшей проработке – более сложные. Сложные модели требуют больших затрат времени на проведение вычислительных экспериментов. Системы уравнений таких моделей обычно отличаются плохой обусловленностью, что создает проблемы обеспечения устойчивости модели, достижения необходимой точности при приемлемых затратах времени.

При поиске оптимального варианта для получения искомого результата приходится многократно проводить эксперимент с учетом многомерности и многокритериальности задачи. На заключительных этапах моделирования часто используют вероятностные модели для того, чтобы исследовать процессы функционирования технической системы в условиях, максимально приближенных к реальным с использованием статистических методов обработки результатов.

Под техническим объектом понимается техническая система – машина, механизм, технический комплекс, технологический процесс, а также любой их компонент, выделяемый в процессе проектирования путем декомпозиции (деления) структуры целостного объекта на отдельные блоки, части, элементы и т. п.

Современная методология моделирования базируется на *системном подходе*. Технический объект при системном подходе рассматривается как сложная система, состоящая из взаимосвязанных, целенаправленно функционирующих элементов и находящаяся во взаимодействии с окружающей внешней средой. Это позволяет учесть все факторы, влияющие на его функционирование, и обеспечить создание технического объекта с высокими показателями эффективности и качества.

1.2. РАЗВИТИЕ ОПРЕДЕЛЕНИЯ СИСТЕМЫ

Термин *система* используют в тех случаях, когда хотят охарактеризовать исследуемый или проектируемый объект как нечто целое (единое), сложное, о котором невозможно сразу дать представление, показав его, изобразив графически или описав математическим выражением (формулой, уравнением и

т. п.). Рассмотрим основные и принципиальные изменения, которые происходили с определением системы по мере развития теории систем и использования этого понятия на практике [7].

В первых определениях говорилось о том, что система – это элементы (части, компоненты) a_i и связи (отношения) r_j между ними:

$$S \underset{def}{\equiv} \langle A, R \rangle, \text{ где } A = \{a_i\}, R = \{r_j\}, \quad (1.1)$$

$$S \underset{def}{\equiv} \langle \{a_i\}, \{r_j\} \rangle, a_i \in A, r_j \in R,$$

$$S \underset{def}{\equiv} [\{a_i\} \& \{r_j\}], a_i \in A, r_j \in R.$$

В приведенных формализованных записях определения использованы различные способы теоретико-множественных представлений: в первых двух – используются различные способы задания множеств и не учитываются взаимоотношения между множествами элементов и связей; в третьем – отражен тот факт, что система – это не простая совокупность элементов и связей того или иного вида, а только те элементы и связи, которые находятся в области пересечения (&) друг с другом (рис. 1.1).

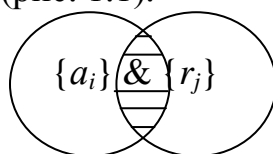


Рис. 1.1. Графическое представление системы, как области пересечения элементов и связей

Так, Л. фон Берталанфи определял систему как *комплекс взаимодействующих компонентов* или как *совокупность элементов, находящихся в определенных отношениях друг с другом и со средой* [7].

В Большой советской энциклопедии система определяется прямым переводом с греческого «συστήμα», что означает «со-став», т. е. *составленное, соединенное из частей*.²

В толковом словаре Ожегова система – это *«нечто целое, представляющее собой единство закономерно расположенных и находящихся во взаимосвязи частей»*.³

Термины «элементы» – «компоненты», «связи» – «отношения» обычно используются (особенно в переводах определений) как синонимы. Однако, строго говоря, «компоненты» – понятие более общее, чем «элементы», может означать совокупность элементов; относительно понятий «связь» и «отношение» существуют разные точки зрения, что будет подробнее рассмотрено далее.

Если известно, что элементы принципиально неоднородны, то это можно сразу учесть в определении, выделив разные множества элементов $A = \{a_i\}$ и $B = \{b_k\}$:

$$S \underset{def}{\equiv} \langle A, B, R \rangle. \quad (1.1a)$$

² БСЭ. Изд. 2-е. Т. 39. С.158.

³ Ожегов С. И. Словарь русского языка. Изд. 8-е. М.: Сов. энц., 1970. 900 с.

В определении М. Месаровича [7], например, выделены множество X входных объектов (воздействующих на систему) и множество Y выходных результатов, а между ними установлено обобщающее отношение пересечения, что можно отобразить либо так же, как у автора определения:

$$S \subseteq X \times Y, \quad S \subseteq X \cap Y, \quad (1.16)$$

либо с помощью других обозначений пересечения:

$$S \subseteq X \bullet Y, \quad S \subseteq X \& Y. \quad (1.1в)$$

Если какой-то вид отношений r_i применяется только к элементам разных множеств и не используется внутри каждого из них, то это можно отразить следующим образом:

$$S \underset{def}{=} \langle a_i, b_k, r_j \rangle, \quad a_i \in A, \quad b_k \in B, \quad r_j \in R, \quad (1.1г)$$

где $\{a_i \ r_i \ b_k\}$ – элементы новой системы, образованные из элементов исходных множеств A и B .

Для уточнения *элементов* и *связей* в определениях включают свойства. Так, в определении А. Холла [7] свойства (атрибуты) Q_A дополняют понятие элемента (предмета):

$$S \underset{def}{=} \langle A, Q_A, R \rangle. \quad (1.1д)$$

А.И. Уёмов, определяя систему через понятия *вещи, свойства, отношения*, предложил двойственные определения [7], в одном из которых свойства q_i характеризуют элементы (*вещи*) a_i , а в другом – свойства q_j характеризуют связи (*отношения*) r_j :

$$S \underset{def}{=} [\{a_i\} \& \{r_j(q_j)\}], \quad a_i \in A, \quad r_j \in R, \quad q_j \in Q_R, \quad (1.1е)$$

$$S \underset{def}{=} [\{a_i(q_i)\} \& \{r_j\}], \quad a_i \in A, \quad q_i \in Q_A, \quad r_j \in R.$$

Далее в определениях системы появляется понятие *цель*. Вначале – в неявном виде: в определении Ф.Е. Темникова [7] *система* – *организованное множество* (в котором *цель* появляется при раскрытии понятия *организованное*); в философском словаре *система* – *совокупность элементов, находящихся в отношениях и связях между собой и образующих некоторое целостное единство*⁴. Потом – в виде конечного результата, системообразующего критерия, функции, а позднее – и с явным упоминанием о цели.

Символически эту группу определений представим следующим образом:

$$S \underset{def}{=} \langle A, R, Z \rangle, \quad (1.2)$$

где Z – *цель*, совокупность или структура целей.

В некоторых определениях уточняются условия целеобразования – *среда* SR , *интервал времени* ΔT , т. е. период, в рамках которого будет существовать система и ее цели, что сделано, например, в определении В.Н. Сагатовского [7]:

⁴ Краткий философский словарь / под ред. М. Розенталя и П. Юдина. Изд. 4-е. М.: Госполитиздат, 1954. 704 с.

система *конечное множество функциональных элементов и отношений между ними, выделенное из среды в соответствии с определенной целью в рамках определенного временного интервала:*

$$S \underset{\text{def}}{=} \langle A, R, Z, SR, \Delta T \rangle. \quad (1.2a)$$

Далее, в определение системы начинают включать, наряду с элементами, связями и целями, наблюдателя N , т. е. лицо, представляющее объект или процесс в виде системы при их исследовании или принятии решения:

$$S \underset{\text{def}}{=} \langle A, R, Z, N \rangle. \quad (1.3)$$

На необходимость учета взаимодействия между изучаемой системой и исследователем указывал еще У.Р. Эшби [7]. Но первое определение, в которое в явном виде включен наблюдатель, дал Ю.И. Черняк: *Система есть отражение в сознании субъекта (исследователя, наблюдателя) свойств объектов и их отношений в решении задачи исследования, познания* [7].

$$S \underset{\text{def}}{=} \langle A, Q_A, R, Z, N \rangle. \quad (1.3a)$$

В последующих вариантах этого определения Ю.И. Черняк стал учитывать и язык наблюдателя L_N , начиная с этого определение: *Система есть отображение на языке наблюдателя (исследователя, конструктора) объектов, отношений и их свойств в решении задачи исследования, познания:*

$$S \underset{\text{def}}{=} \langle A, Q_A, R, Z, N, L_N \rangle. \quad (1.3b)$$

Сопоставляя эволюцию определения системы (*элементы и связи, затем – цель, затем – наблюдатель*) и эволюцию использования категорий теории познания в исследовательской деятельности, можно обнаружить сходство: в начале модели (особенно формальные) базировались на учете только *элементов и связей*, взаимодействий между ними, затем – стало уделяться внимание *цели*, поиску методов ее формализованного представления (целевая функция, критерий функционирования и т. п.), а начиная с 1960-х гг. все большее внимание обращают на *наблюдателя* – лицо, осуществляющее моделирование или проводящее эксперимент, т. е. лицо, принимающее решение.

1.3. ПОНЯТИЯ, ХАРАКТЕРИЗУЮЩИЕ ФУНКЦИОНИРОВАНИЕ СИСТЕМЫ

Обыденная трактовка рассмотренных ниже понятий (элемент, связь и др.) не всегда совпадает с их значением как специальных терминов системного описания и анализа объектов. Поэтому кратко рассмотрим основные понятия, помогающие уточнять представление о системе. Понятия, входящие в определение системы, тесно связаны между собой, и по мнению Л. фон Берталанфи не могут быть определены независимо, а определяются, как правило, одно через другое, уточняя друг друга [7].

Элемент. Под элементом принято понимать простейшую, неделимую часть системы. Однако ответ на вопрос, что является такой частью, может быть неоднозначным.

Например, в качестве элементов стола можно назвать «ножки, ящики, крышку» и т. д., а можно – «атомы, молекулы», в зависимости от того, какая задача стоит перед исследователем.

Поэтому примем следующее определение: *элемент – это предел членения системы с точки зрения аспекта рассмотрения, решения конкретной задачи, поставленной цели.*

Для помощи в выделении элементов при анализе конкретных проблемных ситуаций можно использовать информационный подход и, в частности, меру информации восприятия $J = A/\Delta A$, где ΔA – минимальное количество материального свойства A (квант), с точностью до которого исследователя интересует информация об этом свойстве при формировании модели.

Систему можно расчленять на элементы различными способами в зависимости от формулировки задачи, цели и ее уточнения в процессе проведения системного исследования. При необходимости можно изменять принцип расчленения, выделять другие элементы и получать с помощью нового расчленения более адекватное представление об анализируемом объекте или проблемной ситуации.

Компоненты и подсистемы. Иногда термин элемент используют в более широком смысле, даже в тех случаях, когда система не может быть сразу разделена на составляющие, являющиеся пределом ее членения. Однако при многоуровневом расчленении системы лучше использовать другие термины, предусмотренные в теории систем: сложные системы принято вначале делить на *подсистемы*, или на *компоненты*.

Понятие *подсистема* подразумевает, что выделяется относительно независимая часть системы, обладающая свойствами системы, и в частности, имеющая подцель, на достижение которой ориентирована подсистема, а также другие свойства – свойство целостности, коммуникативности и т. п., определяемые закономерностями систем.

Если же части системы не обладают такими свойствами, а представляют собой просто совокупности однородных элементов, то такие части принято называть *компонентами*.

Расчленяя систему на подсистемы, следует иметь в виду, что так же, как и при расчленении на элементы, выделение подсистем зависит от цели и может меняться по мере ее уточнения и развития представлений исследователя об анализируемом объекте или проблемной ситуации.

Связь. Понятие *связь* входит в любое определение системы и обеспечивает возникновение и сохранение ее целостных свойств. Это понятие одновременно характеризует и строение (статику), и функционирование (динамику) системы.

Связь определяют как ограничение степени свободы элементов. Действительно, элементы, вступая во взаимодействие (связь) друг с другом, утрачивают часть своих свойств, которыми они потенциально обладали в свободном состоянии.

В определениях системы термины *связь* и *отношение* обычно используются как синонимы. Однако существуют разные точки зрения; одни исследователи считают *связь* частным случаем *отношения*; другие – напротив, *отношение* рассматривают

как частный случай *связи*; третьи – предлагают понятие *связь* применять для описания статики системы, ее структуры, а понятием *отношение* характеризовать некоторые действия в процессе функционирования (динамики) системы.

Связи можно охарактеризовать *направлением, силой, характером* (или *видом*). По первому признаку связи делят на *направленные* и *ненаправленные*. По второму – на *сильные* и *слабые* (иногда пытаются ввести «шкалу» силы связей для конкретной задачи). По характеру (виду) различают связи *подчинения*, связи *порождения* (или *генетические*), *равноправные* (или *безразличные*), связи *управления*.

Связи в конкретных системах могут быть одновременно охарактеризованы несколькими из названных признаков.

Важную роль в моделировании систем играет понятие *обратной связи* (рис. 1.2). Обратная связь может быть *положительной*, сохраняющей тенденции происходящих в системе изменений того или иного выходного параметра, и *отрицательной* – противодействующей тенденциям изменения выходного параметра, т.е. направленной на сохранение, стабилизацию требуемого значения параметра (например, стабилизацию выходного напряжения, или в системах организационного управления – количества выпускаемой продукции, ее себестоимости и т. п.).

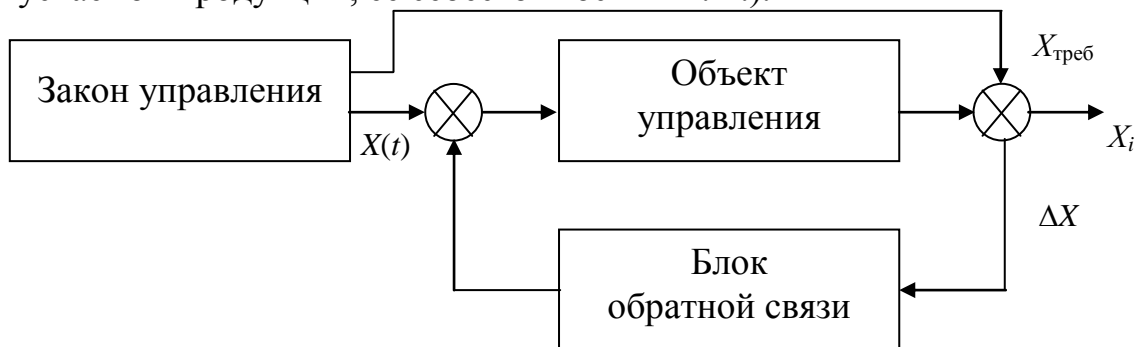


Рис. 1.2. Система управления с обратной связью

При практическом использовании понятия обратной связи часто ограничиваются только фиксацией рассогласования ΔX между требуемым $X_{\text{треб}}$ и фактическим X_i , значением регулируемого параметра, а необходимо учитывать и реализовать все элементы, не забывая замкнуть контур обратной связи, выработав в блоке обратной связи соответствующие управляющие воздействия, которые скорректируют закон управления $X(t)$.

Обратная связь является основой саморегулирования, развития систем, приспособления их к изменяющимся условиям существования.

Цель. Понятие цель и связанные с ним понятия *целесообразности, целенаправленности* лежат в основе развития системы.

Изучению этих понятий большое внимание уделяется в философии, психологии, кибернетике.

Процесс целеобразования и соответствующий ему процесс обоснования целей в организационных системах весьма сложен. На протяжении всего периода развития философии и теории познания происходило развитие представлений о цели.

Анализ определений цели и связанных с ней понятий показывает, что в зависимости от стадии познания объекта, этапа системного анализа, в понятие «цель» вкладывают различные оттенки – от идеальных устремлений (цель – *выражение активности сознания; человек и социальные системы вправе формулировать цели, достижение которых, как им заведомо известно, невозможно, но к которым можно непрерывно приближаться*), до конкретных целей – *конечных результатов*, достижимых в пределах некоторого интервала времени, формулируемых иногда даже в терминах *конечного продукта* деятельности [7].

В некоторых определениях цель как бы трансформируется, принимая различные оттенки в пределах условной «шкалы» – от *идеальных устремлений* к *материальному воплощению, конечному результату деятельности* (рис. 1.3).

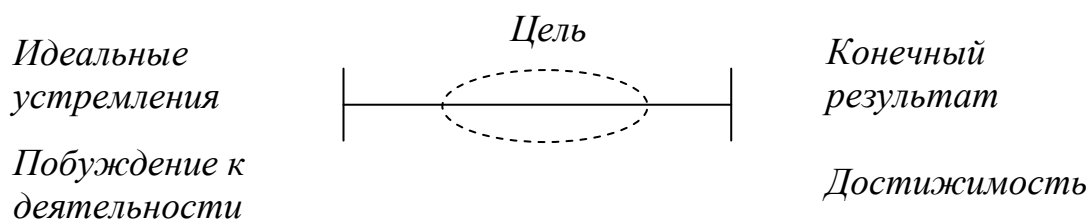


Рис. 1.3. Трансформация определения цели

Например, наряду с приведенным выше определением, целью называется *то, к чему стремится, чему поклоняется и за что борется человек (борется подразумевает достижимость в определенном интервале времени)*; или под целью понимается *модель желаемого будущего* (при этом в понятие «модель» можно вкладывать различные оттенки реализуемости) и, кроме того, вводится понятие, характеризующее разновидность цели (*мечта – это цель, не обеспеченная средствами ее достижения* [7]).

Противоречие, заключенное в понятии «цель», необходимость быть побуждением к действию, *опережающим отражением* (термин введен П.К. Анохиным) или *опережающей идеей*, и одновременно материальным воплощением этой идеи, т. е. быть достижимой, – проявлялось с момента возникновения этого понятия: так, древнеиндийское *артха* означало одновременно *мотив, причину, желание, цель* и даже – *способ* [7].

В русском языке вообще не было термина «цель». Этот термин заимствован из немецкого и имеет значение близкое к понятию «мишень», «финиш», «точка попадания». В английском языке содержится несколько терминов, отражающих различные оттенки понятия цели, в пределах рассматриваемой «шкалы»: *purpose* (цель – намерение, целеустремленность, воля), *object* и *objective* (цель – направление действия, направление движения), *aim* (цель – стремление, прицел, указание), *goal* (цель – место назначения, задача), *target* (цель – мишень для стрельбы, задание, план), *end* (цель – финиш, конец, окончание, предел).

Глубина диалектико-материалистической трактовки понятия цели раскрывается в теории познания, в которой показывается взаимосвязь понятий *цели, оценки, средства, целостности* (и ее *самодвижения*). Изучение

взаимосвязи этих понятий показывает, что в принципе поведение одной и той же системы может быть описано и в терминах цели или целевых функционалов, связывающих цели со средствами их достижения (такое представление называют *аксиологическим*), и без упоминания понятия цели, в терминах непосредственного влияния одних элементов или описывающих их параметров на другие, в терминах «пространства состояний» (или *каузально*). Поэтому одна и та же ситуация в зависимости от склонности и предшествующего опыта исследователя может быть представлена тем или иным способом.

Для того, чтобы отразить диалектическое противоречие, заключенное в понятии «цель», в БСЭ дается следующее определение: *цель – заранее мыслимый результат сознательной деятельности человека, группы людей (заранее мыслимый, но все же результат, воплощение замысла; подчеркивается также, что понятие цели связано с человеком, его сознательной деятельностью, т. е. с наличием сознания, а для характеристики целеустремленных, неэнтропийных тенденций на более низких ступенях развития материи принято использовать другие термины).*

Диалектико-материалистическое понимание цели очень важно при организации процессов коллективного принятия решений в системах управления. В реальных ситуациях необходимо оговаривать, в каком смысле на данном этапе рассмотрения системы используется понятие «цель», что в большей степени должно быть отражено в ее формулировке – *идеальные устремления*, которые помогут коллективу лиц, принимающих решение (ЛПР), увидеть перспективы, или *реальные возможности*, обеспечивающие своевременность завершения очередного этапа на пути к желаемому будущему.

Проведенный анализ определений понятия «цель» и графическая интерпретация «размытости» философских трактовок цели (рис. 1.3), явились важным шагом на пути к практической реализации процессов целеполагания.

Структура. Система может быть представлена простым перечислением элементов или «черным ящиком» (моделью «вход – выход»). Однако чаще всего при исследовании объекта такого представления недостаточного, так как требуется выяснить, что собой представляет объект, что в нем обеспечивает выполнение поставленной цели, получение требуемых результатов. В этих случаях систему отображают путем расчленения на подсистемы, компоненты, элементы с взаимосвязями, которые могут носить различный характер, и вводят понятие структуры.

Структура (от латинского *structure*, означающего строение, расположение, порядок) *отражает определенные взаимосвязи, взаиморасположение составных частей системы, ее устройство (строение).*

При этом в сложных системах структура включает не все элементы и связи, между ними (в предельном случае, когда пытаются применить понятие структуры к простым, полностью детерминированным объектам, понятия структуры и системы совпадают), а лишь наиболее существенные компоненты и связи, которые мало меняются при текущем функционировании системы и обеспечивают существование системы и ее основных свойств. Иными словами, структура характеризует организованность системы, устойчивую упорядоченность элементов и связей.

Структурные связи обладают относительной независимостью от элементов и могут выступать как инвариант при переходе от одной системы к другой, перенося закономерности, выявленные и отраженные в структуре одной из них, на другие. При этом системы могут иметь различную физическую природу.

Одна и та же система может быть представлена разными структурами в зависимости от стадии познания объектов или процессов, от аспекта их рассмотрения, цели создания. При этом по мере развития исследований или в ходе моделирования структура системы может изменяться.

Понятия, характеризующие функционирование и развитие системы.

Процессы, происходящие в сложных системах, как правило, сразу не удастся представить в виде математических соотношений или хотя бы алгоритмов. Поэтому для того, чтобы хоть как-то охарактеризовать стабильную ситуацию или ее изменения, используются специальные термины.

Состояние. Понятием *состояние* обычно характеризуют мгновенную фотографию, «срез» системы, остановку в ее развитии. Его определяют либо через входные воздействия и выходные сигналы (результаты), либо через макропараметры, макросвойства системы (давление, скорость, ускорение). Так, говорят о состоянии покоя (стабильные входные воздействия и выходные сигналы), о состоянии равномерного прямолинейного движения (стабильная скорость) и т. д.

Если рассмотреть элементы h (компоненты, функциональные блоки), учесть, что «входы» можно разделить на управляющие x и возмущающие v (неконтролируемые) и что «выходы» (выходные результаты) зависят от h , x и v , т. е. $y = f(h, x, v)$, то в зависимости от задачи состояние может быть определено как $\{h, x\}$, $\{h, x, y\}$ или $\{h, x, v, y\}$.

Поведение. Если система способна переходить из одного состояния в другое (например, $s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow \dots$), то говорят, что она обладает *поведением*. Этим понятием пользуются, когда неизвестны закономерности (правила) перехода из одного состояния в другое. Тогда говорят, что система обладает каким-то поведением и выясняют его характер, алгоритм.

С учетом введенных обозначений поведение можно представить как функцию $s(t) = [s(t-1), x(t), v(t)]$.

Равновесие. Понятие *равновесие* определяют как способность системы в отсутствии внешних возмущающих воздействий (или при постоянных воздействиях) сохранять свое состояние сколь угодно долго. Это состояние называют *состоянием равновесия*. Простейший пример – равновесие шарика на плоскости.

Устойчивость. Под *устойчивостью* понимают способность системы возвращаться в состояние равновесия после того, как она была из этого состояния выведена под влиянием *внешних* (или в системах с активными элементами – *внутренних*) возмущающих воздействий. Эта способность обычно присуща системам при постоянном u только тогда, когда отклонения не превышают некоторого предела.

Состояние равновесия, в которое система способна возвращаться, называют *устойчивым состоянием равновесия*. Возврат в это состояние может

сопровождаться колебательным процессом. Соответственно в сложных системах возможны неустойчивые состояния равновесия. Простейший пример – устойчивое состояние шарика в ямке до величины отклонений (под воздействием внешних возмущений), которые не выбрасывают его из ямки.

Развитие. Это понятие помогает объяснить сложные термодинамические и информационные процессы. Целесообразно выделять особый класс *развивающихся систем*, обладающих особыми свойствами и требующих разработки и использования специальных подходов к их моделированию.

2. КЛАССИФИКАЦИЯ МОДЕЛЕЙ

2.1. КЛАССИФИКАЦИЯ МОДЕЛЕЙ И МЕТОДОВ ФОРМАЛИЗОВАННОГО ПРЕДСТАВЛЕНИЯ СИСТЕМ

В основу классификации моделей положена степень абстрагирования модели от оригинала [8]. Все модели можно подразделить на две группы: материальные (физические) и абстрактные (математические).

Физической моделью обычно называют систему, которая эквивалентна или подобна оригиналу либо у которой процесс функционирования такой же, как у оригинала, и имеет ту же или другую физическую природу.

Математическая модель представляет собой формализованное описание системы с помощью абстрактного языка, в частности с помощью математических соотношений, отражающих процесс функционирования системы. Для составления модели можно использовать любые математические средства – алгебраическое, дифференциальное и интегральное исчисление, теорию множеств, теорию алгоритмов и т. д. По существу вся математика создана для составления и исследования моделей объектов или процессов.

К средствам абстрактного описания систем относятся также языки химических формул, схем, чертежей, карт, диаграмм и т. п. Выбор вида модели определяется особенностями изучаемой системы и целями моделирования, так как исследование модели позволяет получить ответы на определенную группу вопросов. Для получения другой информации может потребоваться модель другого вида.

Цели моделирования и характерные черты оригинала определяют ряд других особенностей моделей и методы их исследования. Например, математические модели можно классифицировать на детерминированные и вероятностные (стохастические). Первые устанавливают однозначное соответствие между параметрами и характеристиками модели, а вторые – между статистическими значениями этих величин. Выбор того или иного вида модели обусловлен степенью необходимости учета случайных факторов. Среди математических моделей можно выделить по методу их исследования аналитические, численные и имитационные модели.

Аналитической моделью называется такое формализованное описание системы, которое позволяет получить решение уравнения $y = f(h, x, v, t)$ в явном виде, используя известный математический аппарат. *Численная* модель характеризуется зависимостью, которая допускает только частные численные

решения для конкретных начальных условий и количественных параметров модели.

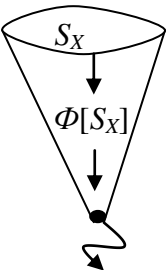
Имитационная модель – это совокупность описания системы и внешних воздействий, алгоритмов функционирования системы или правил изменения состояния системы под влиянием внешних и внутренних возмущений. Эти алгоритмы и правила не дают возможности использования имеющихся математических методов аналитического и численного решения, но позволяют имитировать процесс функционирования системы и производить измерения интересующих характеристик.

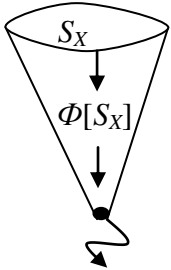
Имитационные модели могут быть созданы для гораздо более широкого класса объектов и процессов, чем аналитические и численные модели. Поскольку для реализации имитационных моделей используются вычислительные системы, средствами формализованного описания имитационных моделей служат универсальные или специальные языки имитационного моделирования.

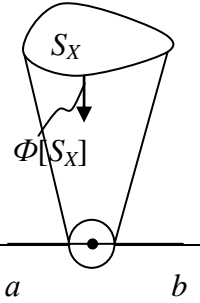
В табл. 2.1 представлена укрупненная классификация групп-направлений методов формализованного представления систем [7]. Эти направления непрерывно развиваются.

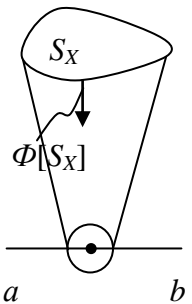
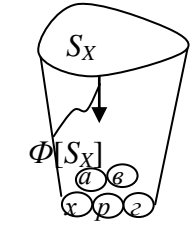
Таблица 2.1

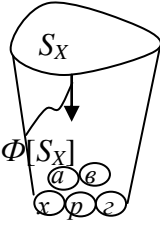
Классификация методов формализованного представления систем

Название класса методов и символический «образ»	Основная терминология и примеры теорий, возникших и развивающихся на базе соответствующего класса методов	Сфера и возможности применения
1	2	3
<p>Аналитические методы</p> 	<p><i>Аналитическими</i> названы здесь методы, в которых ряд свойств многомерной, многосвязной системы (или какой-либо ее части) отображается в n-мерном пространстве одной единственной точкой, совершающей какое-либо движение. Это отображение осуществляется либо с помощью функции $f(S_X)$, либо посредством оператора (функционала) $\Phi(S_X)$. Можно также две или более системы либо их части</p>	<p>Применяются в тех случаях, когда свойства системы можно отобразить с помощью детерминированных величин или зависимостей, т. е. когда знания о процессах и событиях в некотором интервале времени позволяют полностью определить поведение их вне этого интервала. Такие методы используются при решении задач движения и устойчивости, оптимального размещения, распределения работ и ресурсов, выбора наилучшего пути,</p>

1	2	3
	<p>отобразить точками и рассматривать взаимодействие этих точек, каждая из которых совершает какое-то движение, имеет свое поведение. Поведение точек и их взаимодействие описываются аналитическими закономерностями. Основу понятийного (терминологического) аппарата составляют понятия классической математики (<i>величина, функция, уравнение, система уравнений</i> и т. п.). На базе аналитических представлений возникли и развиваются математические теории различной сложности – от аппарата классического математического анализа (методов исследования экстремумов функций, вариационного исчисления и т. п.) до таких разделов современной математики, как <i>математическое программирование</i> (линейное, нелинейное, динамическое и т. п.), <i>теория игр</i> (матричные игры с чистыми стратегиями, дифференциальные игры)</p>	<p>оптимальной стратегии поведения в конфликтных ситуациях и т. п. Математические теории, развивающиеся на базе аналитических представлений, явились основой ряда прикладных теорий (теории автоматического управления, теории оптимальных решений и др.). При практическом применении аналитических представлений для отображения сложных систем следует иметь в виду, что они требуют установления всех детерминированных взаимосвязей между учитываемыми компонентами и целями моделирования системы в виде аналитических зависимостей. Для сложных многокомпонентных, многокритериальных систем получить требуемые аналитические зависимости очень трудно. Более того, если даже это и удастся, то практически невозможно доказать правомерность применения этих аналитических выражений т. е. адекватность модели рассматриваемой задаче</p>

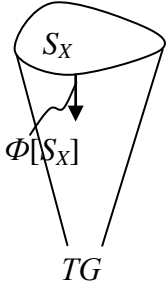
1	2	3
<p>Статистические методы</p> 	<p>В тех случаях, когда не удастся представить систему с помощью детерминированных категорий, можно применить отображение ее с помощью случайных (<i>стохастических</i>) событий, процессов, которые описываются соответствующими вероятностными (<i>статическими</i>) характеристиками и <i>статическими закономерностями</i>. <i>Статистические</i> отображения системы в общем случае (по аналогии с аналитическими) можно представить как бы в виде «размытой» точки (размытой области) в n-мерном пространстве, в которую переводит систему (ее учитываемые свойства) оператор $\Phi(S_X)$. «Размытую» точку следует понимать как некоторую область, характеризующую движение системы (ее поведение); при этом границы области заданы с некоторой вероятностью («размыты») и движение точки определяется некоторой случайной функцией. Закрепляя все параметры, кроме одного, можно получить срез по линии $a - b$, физический смысл которого – воздействие данного параметра на поведение системы, что можно описать статистическим распределением по этому</p>	<p>Статистические отображения позволили расширить области применения ряда дисциплин, возникших на базе аналитических представлений. Так возникли <i>статистическая теория распознавания образов, стохастическое программирование</i>, новые разделы теории игр и др. Расширение возможностей отображения сложных систем и процессов по сравнению с аналитическими методами можно объяснить тем, что при применении статистических представлений процесс постановки задачи как бы частично заменяется статистическими исследованиями, позволяющими, не выявляя все детерминированные связи между изучаемыми событиями или учитываемыми компонентами сложной системы, на основе выборочного исследования (исследования <i>представительной выборки</i>) получать статистические закономерности и распространять их на поведение системы в целом. Однако не всегда можно получить статистические закономерности, не всегда может быть определена <i>представительная</i> (репрезентативная) <i>выборка</i>, доказана правомерность статистических</p>

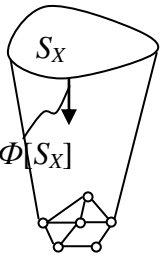
1	2	3
	<p>параметру. Аналогично можно получить двумерную, трехмерную и т. д. картины статистического распределения. На статистических отображениях базируются теории <i>математической статистики, теории статистических испытаний или статистического имитационного моделирования</i> (частным случаем которой является метод Монте-Карло), <i>теория выдвижения и проверки статистических гипотез</i> (частным случаем которого является байесовский подход к исследованию процессов передачи информации в процессах общения, обучения и других ситуациях, характерных для сложных развивающихся систем)</p>	<p>закономерностей. В ряде случаев для получения статистических закономерностей требуются недопустимо большие затраты времени, что также ограничивает возможности их применения</p>
<p>Теоретико-множественные представления</p> 	<p><i>Теоретико-множественные представления, предложенные Г. Кантором, базируются на понятиях: множество (содержательно эквивалентное понятиям «совокупность», «собрание», «ансамбль», «коллекция» и т. п.), элементы множества и отношения на множествах.</i> Сложную систему можно отобразить в виде совокупности разнородных множеств и отношений между ними. Множества могут задаваться двумя способами; <i>перечислением элементов (a_1, a_2, \dots, a_n) и названием</i></p>	<p>Благодаря тому, что при теоретико-множественных представлениях систем и процессов в них можно вводить любые отношения, эти представления: а) служат хорошим языком, с помощью которого облегчается взаимопонимание между представителями различных областей знаний; б) могут являться основой для возникновения новых научных направлений, для создания <i>языков моделирования, языков автоматизации проектирования</i>. Теоретико-множественные представления</p>

1	2	3
	<p><i>характеристического свойства</i> (именем, отражающим это свойство) – например, множество A. В основе большинства теоретико-множественных преобразований лежит переход от одного способа задания множества к другому. В множестве могут быть выделены подмножества. Из двух и более множеств или подмножеств можно, установив отношения между их элементами, сформировать новое множество, состоящее из элементов, качественно отличающихся от элементов исходных множеств (при таком преобразовании у элементов нового множества как бы появляется иной смысл по сравнению с исходными). Теоретико-множественные представления допускают введение любых отношений. При конкретизации применяемых отношений и правил их использования можно получить одну из алгебр логики, один из формальных языков математической лингвистики, создать язык моделирования сложных систем, который затем, получив соответствующее название, может развиваться как самостоятельное научное направление</p>	<p>являются основой математической теории систем М. Месаровича. Однако свобода введения любых отношений приводит к тому, что в создаваемых языках моделирования трудно ввести правила, закономерности, используя которые формально, можно получить новые результаты, адекватные реальным моделируемым объектам и процессам (как это позволяют делать аналитические и статические методы). Поэтому первоначально при применении теоретико-множественных представлений стремились использовать ограниченный набор отношений. В общем же случае в языке могут появляться ситуации <i>парадоксов</i> или <i>антиномий</i>, что приводит к необходимости ограничения разнообразия отношений в создаваемых языках</p>

1	2	3
<p>Логические методы</p> 	<p><i>Логические</i> представления переводят реальную систему и отношения в ней на язык одной из <i>алгебр логики</i> (двузначной, многозначной), основанных на применении алгебраических методов для выражения законов формальной логики. Наибольшее распространение получила бинарная алгебра логики Буля (булева алгебра). Алгебра логики оперирует понятиями: высказывание, <i>предикат</i>, логические операции (<i>логические функции, кванторы</i>). В ней доказываются теоремы, приобретающие затем силу <i>логических законов</i>, применяя которые, можно преобразовать систему из одного описания в другое с целью ее совершенствования, например получить более простую структуру (схему), содержащую меньшее число состояний, элементов, но осуществляющую требуемые функции. Теоремы доказываются и используются в рамках формального <i>логического базиса</i>, который определяется совокупностью специальных правил. Логические методы представления систем относятся к детерминистическим, хотя возможно и их расширение в сторону вероятностных оценок.</p>	<p>Применяются при исследовании новых структур систем разнообразной природы (технических объектов, текстов и др.), в которых характер взаимодействия между элементами еще не настолько ясен, чтобы было возможно их представление аналитическими методами, а статистические исследования либо затруднены, либо не привели к выявлению устойчивых закономерностей. В то же время следует иметь в виду, что с помощью логических алгоритмов можно описывать не любые отношения, а лишь те, которые предусмотрены законами алгебры логики и подчиняются требованиям логического базиса. Логические представления нашли широкое практическое применение при исследовании и разработке автоматов разного рода, автоматических систем контроля, а также при решении задач распознавания образов. Логические представления лежат в основе <i>теории алгоритмов</i>. На их базе развиваются прикладные разделы теории формальных языков. В то же время смысловыражающие возможности логических методов ограничены базисом и функциями алгебры логики и не всегда позволяют адекватно</p>

1	2	3
	<p>На базе математической логики созданы и развиваются <i>теории логического анализа и синтеза, теория автоматов</i>. На основе логических представлений первоначально начинали развиваться некоторые разделы <i>теории формальных языков</i>. В силу ограниченности смысловыражающих возможностей бинарной логики алгебры логики в последнее время имеются попытки создания многозначных (тернарной и т. п.) алгебр логики с соответствующими логическими базисами и теоремами</p>	<p>отобразить реальную проблемную ситуацию. Попытки же создания многозначных алгебр логики на практике пока не находят широкого применения из-за сложности создания логического базиса и доказательства формальных теорем-законов многозначной алгебры логики</p>
<p>Лингвистические, семиотические представления</p> 	<p><i>Лингвистические</i> представления базируются на понятиях <i>тезауруса T</i> (множества смысловыражающих элементов языка с заданными смысловыми отношениями; он характеризует структуру языка), <i>грамматики G</i> (правил образования смысловыражающих элементов разных уровней тезауруса), <i>семантики</i> (смыслового содержания формируемых фраз, предложений и других смысловыражающих элементов) и <i>прагматики</i> (смысла для заданной задачи, цели).</p>	<p>Лингвистические и семиотические представления возникли и развиваются в связи с потребностями анализа текстов и языков. Однако в последнее время эти представления начинают широко применяться для отображения и анализа процессов в сложных системах в тех случаях, когда не удастся применить сразу аналитические, статистические представления или методы формальной логики. В частности, лингвистические и семиотические представления являются удобным аппаратом (особенно о сочетании с графическими) для первого этапа постепенной формализации задач принятия решений в плохо</p>

1	2	3
	<p>Семиотические представления базируются на понятиях: знак, знаковая система, знаковая ситуация. Семиотика возникла как наука о знаках в широком смысле. Однако наиболее широкое практическое применение нашло направление лингвистической семиотики, которое наряду с основными понятиями семиотики (знак, знаковая система, треугольник Фреге и т. п.) широко пользуется некоторыми понятиями математической лингвистики (тезаурус, грамматика и т. п.). С теоретической точки зрения границу между лингвистическими и семиотическими представлениями при разработке языков моделирования можно определить характером правил грамматики (если правила не охватываются классификацией правил вывода формальных грамматик Н. Хомского, то модель удобнее отнести к семиотической и применять принципы её анализа, предлагаемые семиотикой). Для практических приложений модели лингвистических и семиотических представлений можно рассматривать как один класс методов формализованного представления систем</p>	<p>формализуемых ситуациях, чем и был вызван возрастающий интерес к этим методам со стороны инженеров и разработчиков сложных систем. На их основе разрабатывают языки моделирования, автоматизации проектирования и т. д. Что касается недостатков методов, то при усложнении языка моделирования, при применении правил произвольных грамматик Н. Хомского или правил лингвистической семиотики трудно гарантировать правильность получаемых результатов, возникают проблемы алгоритмической разрешимости. Возможно появление парадоксов, что частично может быть устранено с помощью содержательного контроля и корректировки языка на каждом шаге его расширения в диалоговом режиме моделирования. При этом создатель языка не всегда может объяснить его возможности, происходит как бы «выращивание» языка, у которого появляются новые свойства</p>

1	2	3
<p>Графические представления</p> 	<p>К <i>графическим</i> представлениям здесь отнесены любые графики (<i>графики Ганта, диаграммы, гистограммы</i> и т. п.) и возникшие на основе графических отображений теории; <i>теория графов, теория сетевого планирования и управления</i> и т. п.), т. е. все то, что позволяет наглядно представить процессы, происходящие в системах, и облегчить, таким образом, их анализ для человека (лица, принимающего решения)</p>	<p>Графические представления являются удобным средством исследования структур и процессов в сложных системах и решения различного рода организационных вопросов в информационно-управляющих комплексах, в которых необходимо взаимодействие человека и технических устройств.</p> <p>Широкое применение на практике получила теория сетевого планирования и управления</p>

Рассмотрим первые два метода (аналитические и статистические) более подробно и кратко охарактеризуем их особенности, достоинства и недостатки с точки зрения возможности использования при моделировании систем. Эти группы методов получили наибольшее распространение в практике моделирования и управления. Для представления промежуточных и окончательных результатов моделирования широко используются графические представления (графики, диаграммы и т. п.). Однако последние являются вспомогательными; основу же модели, доказательства ее адекватности составляют те или иные направления аналитических и статистических представлений.

Аналитическими в рассматриваемой классификации названы методы, которые отображают реальные объекты и процессы в виде точек (безразмерных в строгих математических доказательствах), совершающих какие-либо перемещения в пространстве или взаимодействующих между собой.

В табл. 2.1 эта особенность аналитических представлений условно иллюстрируется символическим образом – преобразованием сложной системы в точку, совершающую какое-то движение (или обладающую каким-то поведением), посредством оператора (функции, функционала) $\Phi(S_X)$. Как правило, поведение точек и их взаимодействие описываются строгими соотношениями, имеющими силу *закона*.

Основу понятийного (терминологического) аппарата этих представлений составляют понятия классической математики (*величина, формула, функция, уравнение, система уравнений, логарифм, дифференциал, интеграл* и т. д.).

Аналитические представления имеют многовековую историю развития и для них характерно не только стремление к строгости терминологии, но и к закреплению за некоторыми специальными величинами определенных букв

(например, удвоенное отношение площади круга к площади вписанного в него квадрата – $\pi = 3,14$; основание натурального логарифма $e = 2,7$ и т. д.).

На базе аналитических представлений возникли и развиваются математические теории различной сложности – от аппарата классического *математического анализа* (методов исследования функций, их вида, способов представления, поиска экстремумов функций и т. п.) до таких новых разделов современной математики, как *математическое программирование* (линейное, нелинейное, динамическое и т. п.), *теория игр* (матричные игры с чистыми стратегиями, дифференциальные игры и т. п.).

Эти теоретические направления стали основой многих прикладных дисциплин, в т. ч. теории автоматического управления, теории оптимальных решений и т. д.

При моделировании систем применяется широкий спектр символических представлений, использующих «язык» классической математики. Однако далеко не всегда эти символические представления адекватно отражают реальные сложные процессы, и их в этих случаях, вообще говоря, нельзя считать строгими математическими моделями.

Большинство из направлений математики не содержит средств постановки задачи и доказательства адекватности модели. Последняя доказывается экспериментом, который по мере усложнения проблем становится также все более сложным, дорогостоящим, не всегда бесспорен и реализуем.

В то же время в состав этого класса методов входит такое направление математики, как *математическое программирование*, которое содержит средства постановки задачи и расширяет возможности доказательства адекватности моделей.

Идея этого направления была предложена инженером, а впоследствии за работы в этой области – лауреатом Государственной и Нобелевской премий Л.В. Канторовичем [9] для решения экономических задач (в частности, задачи раскроя фанеры).

Привлекательность методов математического программирования для решения слабоформализованных задач (каковыми, как правило, являются задачи планирования, распределения работ и ресурсов, загрузки оборудования и другие задачи управления современным предприятием на начальном этапе их постановки) объясняется рядом особенностей, отличающих эти методы от методов классической математики.

Для пояснения этих особенностей рассмотрим упрощенный пример.

Предположим, что в трех цехах $Ц1$, $Ц2$, $Ц3$ изготавливаются два вида изделий $И1$ и $И2$. Известна удельная загрузка каждого цеха a_j (оцениваемая в данном случае в процентах) при изготовлении каждого из изделий и прибыль (или цена, объем реализуемой продукции в рублях) c_i от реализации изделий. Требуется определить, сколько изделий каждого вида следует производить при возможно более полной загрузке цехов, чтобы получить за рассматриваемый плановый период максимальную прибыль или максимальный объем реализуемой продукции.

Условия задачи удобно представить в форме табл. 2.2, которая характерна

для задач математического программирования, т. е. целевой функции (в данном случае определяющей максимизацию прибыли)

$$F = \sum_{i=1}^2 c_i x_i = 240x_1 + 320x_2 \rightarrow \max = v \quad (2.1)$$

с рядом ограничений (в данном случае диктуемых возможностями цехов, т. е. их предельной 100%-ной загрузкой)

$$\begin{cases} 5x_1 + 4x_2 \leq 100, \\ 1,6x_1 + 6,4x_2 \leq 100, \\ 2,9x_1 + 5,8x_2 \leq 100. \end{cases} \quad (2.2)$$

Таблица 2.2

Удельная загрузка цехов, a_{ij}

Изделия, i	Цех j , %			Цена изделия c_i , руб.
	Ц1	Ц2	Ц3	
И1	5	1,6	2,9	240
И2	4	6,4	5,8	320
Максимальная загрузка, B_j	100	100	100	$v = 240$

В данном случае ограничения однородны и их можно записать короче:

$$\sum_{i=1}^2 a_{ij} < B_j, \quad j = \overline{1,3}. \quad (2.2a)$$

В общем случае может быть несколько групп подобных ограничений (например, по имеющимся материалам разного вида, себестоимости, заработной плате рабочих и т. п.).

Графическое решение задачи приведено на рис. 2.1. Ограничения определяют область допустимых решений, а наклон прямой, отображающей целевую функцию, определяет точку последнего ее пересечения с областью допустимых решений, которая и является наилучшим решением задачи (экстремумом или оптимумом). В данном случае $x_1 = 9$, $x_2 = 13$.

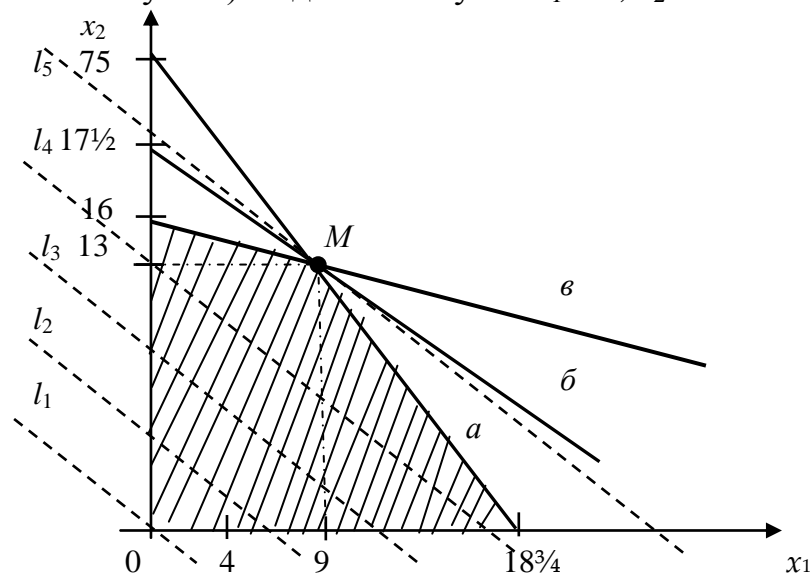


Рис. 2.1. Графическое решение задачи

В случае большего числа разнородных ограничений графическая интерпретация задачи затруднена, поэтому используются специальные методы (например, симплекс-метод) или пакеты прикладных программ, их реализующие. В зависимости от вида целевой функции ограничений и принципов организации решения выделяют направления математического программирования: *линейное* (при линейном характере целевой функции ограничений), *нелинейное* (целевая функция нелинейна), *целочисленное* (ограничение на характер переменных), *динамическое* и т. п. Эти направления имеют специфические особенности и методы решения. Но основная суть постановки задачи сохраняется.

Анализ хода постановки и решения задачи позволяет выявить основные особенности математического программирования:

- введение понятий *целевая функция* и *ограничения* и ориентация на их формирование является фактически некоторыми *средствами постановки задачи*. Эти средства могут использоваться даже если не удастся сформировать систему непротиворечивых ограничений или записать целевую функцию в формальном виде. В итоге удастся уточнить представление о проблемной ситуации и, таким образом, поставить задачу хотя бы в первом приближении;
- при использовании методов математического программирования появляется *возможность объединения* в единой модели *разнородных критериев* (разных размерностей, предельных значений), что очень важно при анализе реальных проектных и производственных ситуаций;
- модель математического программирования допускает (и даже ориентирует на это) *выход на границу области определения переменных*;
- изучение методов решения задач математического программирования позволяет получить представление о пошаговом приближении к решению, т. е. о *пошаговом алгоритме* получения результата моделирования;
- графическая интерпретация задачи дает наглядное представление об области допустимых решений (на рис. 2.1 заштрихована), что помогает в практических ситуациях даже в тех случаях, когда не удастся получить формальное отображение целевой функции и строго решить задачу математического программирования.

Благодаря рассмотренным особенностям методы математического программирования можно кратко охарактеризовать как методы, имеющие в отличие от классической математики некоторые средства постановки задачи. В частности, термин *целевая функция* часто используется даже в тех случаях, когда очевидна невозможность формального установления детерминированных взаимосвязей между компонентами и целями системы. Помогает в постановке задачи и понятие *области допустимых решений*. Этим объясняется популярность рассматриваемого направления. Однако получаемые в таких случаях модели уже не будут относиться к моделям математического программирования и к аналитическим методам.

Аналитические методы применяются в тех случаях, когда свойства

системы можно отобразить с помощью детерминированных величин или зависимостей, т. е. когда знания *о процессах и событиях в некотором интервале времени позволяют полностью определить поведение их вне этого интервала*. Эти методы используются при решении задач движения и устойчивости, оптимального размещения, распределения работ и ресурсов, выбора наилучшего пути, оптимальной стратегии поведения, в т. ч. в конфликтных ситуациях и т. п.

В то же время при практическом применении аналитических представлений для отображения сложных систем следует иметь в виду, что они требуют установления всех детерминированных связей между учитываемыми компонентами и целями системы в виде аналитических зависимостей. Для сложных многокомпонентных, многокритериальных систем получить требуемые аналитические зависимости крайне трудно. Более того, даже если это и удастся, то практически невозможно доказать *правомерность применения таких выражений*, т. е. *адекватность модели рассматриваемой задаче*.

В таких ситуациях следует обратиться к другим методам моделирования.

Статистические представления сформировались как самостоятельное научное направление в середине прошлого века (хотя возникли значительно раньше). Основу их составляет отображение явлений и процессов с помощью случайных (*стохастических*) событий и их поведений, которые описываются соответствующими вероятностными (*статистическими*) характеристиками и *статистическими закономерностями*.

Термин *стохастические* уточняет понятие «случайный», которое в обыденном смысле принято связывать с отсутствием причин появления событий, с появлением не только повторяющихся и подчиняющихся каким-то закономерностям, но и единичных событий. Процессы же, отображаемые статистическими закономерностями, должны быть жестко связаны с заранее заданными, определенными причинами, а «случайность» означает, что они могут появиться или не появиться при наличии заданного комплекса причин.

Статистические отображения системы в общем случае (по аналогии с аналитическими) можно представить (см. символический образ в табл. 2.1) как бы в виде «размытой» точки (размытой области) в n -мерном пространстве, в которую переводит систему (ее учитываемые в модели свойства) оператор $\Phi(S_X)$. «Размытую» точку следует понимать как некоторую область, характеризующую движение системы (ее поведение); при этом границы области заданы с некоторой вероятностью p («размыты») и движение точки описывается некоторой случайной функцией.

Закрепляя все параметры этой области, кроме одного, можно получить срез по линии $a - b$, смысл которого – воздействие данного параметра на поведение системы, что можно описать статистическим распределением по этому параметру. Аналогично можно получить двумерную, трехмерную и т. д. картины статистического распределения.

Статистические закономерности можно представить или в виде *дискретных* случайных величин и их вероятностей, или в виде *непрерывных* зависимостей распределения событий, процессов.

Для *дискретных* событий соотношение между возможными значениями x_i случайной величины X и их вероятностями $p_i = p(x_i) = P(X = x_i)$ называют *законом распределения* и записывают в виде ряда (табл. 2.3) либо представляют в виде зависимостей $F(x)$ (рис. 2.2, а) или

$$p(x) = \begin{cases} p_i, & \text{если } x = x_i, \\ 0, & \text{иначе если } x = x_i \end{cases} \quad (\text{рис. 2.2, в}). \text{ При этом}$$

$$F(x) = \sum_{x_i < x} p(x_i). \quad (2.3)$$

Таблица 2.3

x	x_1	x_1	...	x_i	...	x_n
$p(x)$	p_1	p_2	...	p_i	...	p_n

Для *непрерывных* случайных величин (процессов) *закон распределения* представляют (в соответствии с дискретными законами) либо в виде *функции распределения* (*интегральный закон распределения* – рис. 2.2, б), либо в виде *плотности вероятностей* (*дифференциальный закон распределения* – рис. 2.2, г). В этом случае $p(x) = dF(x)/dx$ и $\Delta F(x) = p(x)\Delta x$, где $p(x)$ – вероятность попадания случайных событий в интервал от x до $x + \Delta x$.

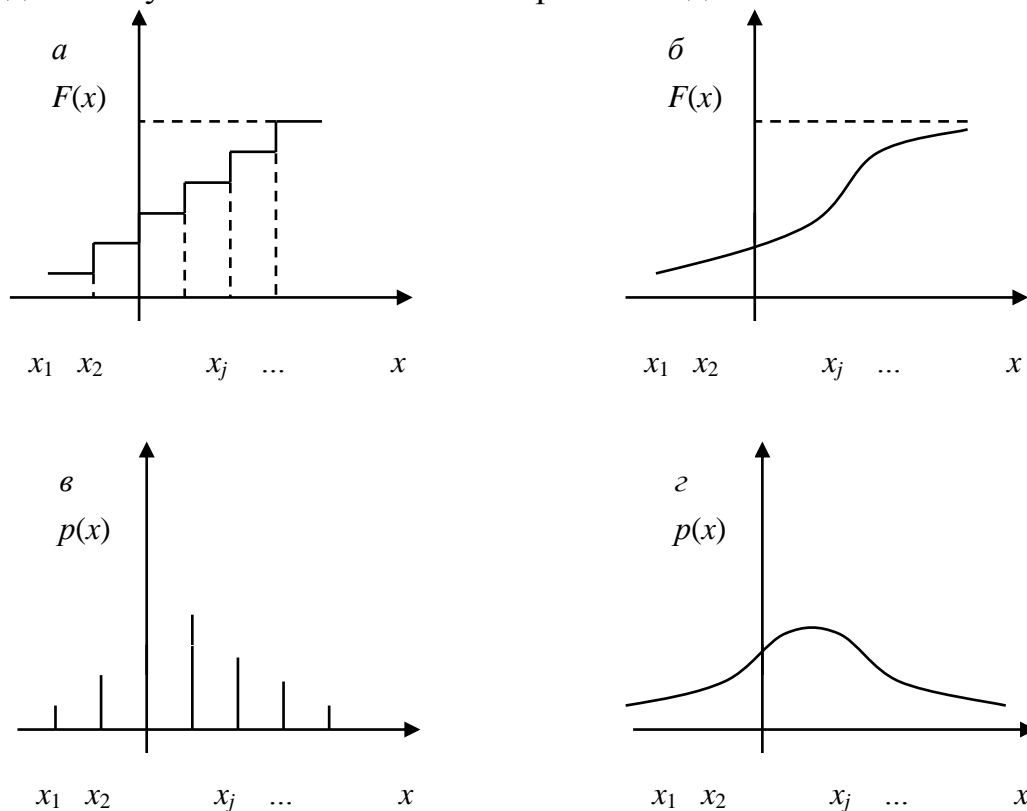


Рис. 2.2. Законы распределения для дискретных событий (а, в) и непрерывных случайных величин (б, г)

Для полной группы несовместных событий имеет место условие нормирования. Сумма их вероятностей равна 1. Для функциональных характеристик это условие имеет вид

$$F(\infty) = \sum_{i=1}^n p = 1. \quad (2.4)$$

$$\int_{-\infty}^{\infty} p(x)dx = F(\infty) - F(-\infty) = 1 - 0 = 1. \quad (2.4a)$$

На практике применяют тот или иной вид зависимостей, приведенных на рис. 2.2, более подходящий для соответствующих приложений.

Закон распределения является удобной формой статистического отображения системы. Однако получение закона (даже одномерного) или определение изменений этого закона при прохождении через какие-либо устройства или среды представляет собой трудную, часто невыполнимую задачу. Поэтому в ряде случаев пользуются не распределением, а его характеристиками – начальными центральными моментами.

Наибольшее применение получили:

- 1-й начальный момент – *математическое ожидание* или *среднее значение* случайной величины

$$\begin{aligned} \bar{x} = \mu_x &= \sum_{i=1}^n x_i p_i \text{ – для дискретных величин;} \\ \bar{x} = \mu_x &= \int_{-\infty}^{\infty} p(x)dx \text{ – для непрерывных величин;} \end{aligned} \quad (2.5)$$

- 2-й центральный момент – *дисперсия* случайной величины:

$$\begin{aligned} \sigma_x^2 &= \sum_{i=1}^n (x_i - \mu_x)^2 p_i \text{ – для дискретных величин;} \\ \sigma_x^2 &= \int_{-\infty}^{\infty} (x - \mu_x)^2 p(x)dx \text{ – для непрерывных величин.} \end{aligned} \quad (2.6)$$

Если дискретная случайная величина X задана некоторой выборкой $\{x_1, \dots, x_n\}$, то несмещенная оценка дисперсии равна

$$\sigma_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu_x)^2 p_i. \quad (2.6a)$$

На практике иногда используется не дисперсия σ_x^2 , а *среднее квадратическое отклонение* σ_x .

Связь между двумя случайными величинами в общем случае характеризуется *ковариацией* – моментом связи:

$$\text{cov}(x, y) = \mu_{xy} = M[(x - \mu_x)(y - \mu_y)]. \quad (2.7)$$

Ковариация нормированных отклонений – *коэффициент корреляции*

$$r_{xy} = \text{cov}(x', y') = \frac{M(x - \mu_x)(y - \mu_y)}{\sigma_x \sigma_y}, \quad (2.8)$$

где $x' = (x - \mu_x)/\sigma_x$, $y' = (y - \mu_y)/\sigma_y$ – нормированные отклонения; σ_x , σ_y – средние квадратические отклонения.

В случае, когда $\{x_i, y_i\}_{i=1}^n$ является выборкой, коэффициент корреляции

вычисляется по формуле

$$r_{xy} = \frac{1}{n-1} \sum_{i=1}^n \left[\frac{(x_i - \mu_x)(y_i - \mu_y)}{\sigma_x \sigma_y} \right]. \quad (2.8a)$$

Практическое применение получили в основном одномерные распределения, что связано со сложностью получения статистических закономерностей и доказательства адекватности применения двумерных величин для конкретных приложений. При этом, как правило, используется понятие *выборки*.

Под *выборкой* понимается часть изучаемой совокупности явлений, на основе исследования которой получают статистические закономерности, присущие всей совокупности и распространяемые на нее с какой-то вероятностью.

Для того чтобы полученные при исследовании выборки закономерности можно было распространить на всю совокупность, выборка должна быть *представительной* (*репрезентативной*), т. е. обладать определенными качественными и количественными характеристиками. Качественные характеристики связаны с содержательным аспектом выборки, т. е. с определением, являются ли элементы, входящие в выборку, элементами исследуемой совокупности, правильно ли отобраны эти элементы с точки зрения цели исследования (с этой точки зрения выборка может быть *случайной*, *направленной* или *смешанной*). Количественные характеристики представительности выборки связаны с определением объема выборки, достаточного для того, чтобы на основе ее исследования можно было делать выводы о совокупности в целом. В большинстве практических случаев вопрос о количественных характеристиках выборки является предметом специального исследования.

На базе статистических представлений развивается ряд математических теорий: *математическая статистика*, объединяющая различные методы статистического анализа (регрессионный, дисперсионный, корреляционный, факторный и т. п.); *теория статистических испытаний*, основой которой является метод Монте-Карло, а развитием – *теория статистического имитационного моделирования*; *теория выдвижения и проверки статистических гипотез*, возникшая для оценки процессов передачи сигналов на расстоянии и базирующаяся на общей *теории статистических решающих функций* А. Вальда. Частным случаем теории выдвижения гипотез, важным для теории систем, является *байесовский подход* к исследованию процессов передачи информации в процессах общения, обучения и других ситуациях. В организационных системах используется *теория потенциальной помехоустойчивости*, начала которой положены в работах В.А. Котельникова. Последние два направления обобщает *теория статистических решений*, в рамках которой, в свою очередь, возник ряд интересных и полезных для практики направлений.

Перечисленные направления в большинстве своем носят теоретико-прикладной характер и возникли из потребностей практики. Однако есть и ряд

дисциплин, которые носят более выраженный прикладной характер. В их числе *статистическая радиотехника, статистическая теория распознавания образов, экономическая статистика, теория массового обслуживания*, а также развившиеся из направлений, возникших на базе аналитических представлений, *стохастическое программирование*, новые разделы *теории игр* и т. п.

Расширение возможностей отображения сложных систем и процессов по сравнению с аналитическими методами можно объяснить тем, что при применении статистических представлений процесс постановки задачи как бы частично заменяется статистическими исследованиями. Последние позволяют, не выявляя все детерминированные связи между изучаемыми объектами или компонентами сложной системы, на основе выборочного исследования получать статистические закономерности и распространять их на поведение системы в целом.

Однако не всегда можно получить статистические закономерности, не всегда может быть определена репрезентативная выборка, доказана правомерность применения статистических закономерностей. Если же не удастся доказать репрезентативность выборки или для этого требуется недопустимо большое время, то применение статистических методов может привести к неверным результатам.

2.2. КЛАССИФИКАЦИЯ ВЕРОЯТНОСТНЫХ МОДЕЛЕЙ

Вероятностные (стохастические) модели описывают ситуации, в которых похожие причины приводят к различным следствиям, т. е. имеет место элемент случайности. Для построения вероятностной модели необходимо знать, какие величины можно считать случайными, а какие – неслучайными; какой характер имеют законы распределения случайных величин и т. д.

Вероятностные модели можно разделить на две большие группы (рис. 2.3) [7]:

- математическая модель, в которой можно точно указать законы распределения случайных величин, является *теоретико-вероятностной*;
- математическая модель, в которой заранее нельзя указать законы распределения случайных величин, является *статистической*.

По степени сложности вероятностные модели делятся на три уровня. Простейшие теоретико-вероятностные модели первого уровня – *случайное событие* (СС) и *случайная величина* (СВ), являющиеся соответственно качественной и количественной характеристиками проведенного испытания.

СС может быть простейшим (элементарным) или сложным (выраженным через элементарные). Для описания вероятностных свойств простейшего случайного события A используется стандартная формула классической вероятности:

$$P(A) = \frac{m}{n},$$

где m – число случаев из пространства элементарных событий, благоприятных событию A ; n – общее число случаев, содержащееся в пространстве элементарных событий.

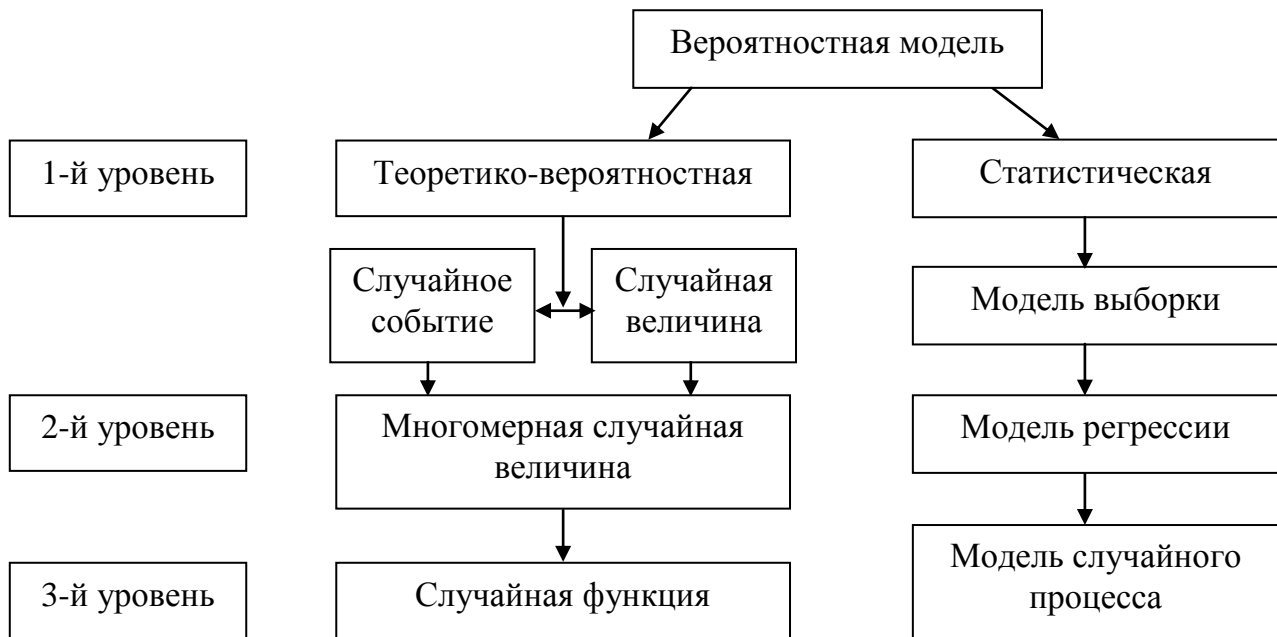


Рис. 2.3. Классификация вероятностных моделей

Геометрическая форма классической формулы имеет вид

$$P(A) = \frac{\mu(A)}{\mu(D)},$$

где $\mu(A)$ – мера области благоприятных случаев A ; $\mu(D)$ – мера области всех случаев D . При этом $A \subseteq D$.

Для расчета вероятности используются теоремы сложения и умножения вероятностей:

$$P(A + B) = P(A) + P(B) - P(A \cdot B);$$

$$P(A \cdot B) = P(A/B) \cdot P(B),$$

где $P(A/B)$ – условная вероятность события A при условии, что событие B произошло.

Пусть $\{H_i\}_{i=1}^n$ – полная несовместная группа событий (гипотез), тогда для определения вероятности сложных событий используется формула полной вероятности:

$$P(A) = \sum_{i=1}^n P(H_i) \cdot P(A/H_i),$$

где $P(H_i)$ – вероятность гипотезы H_i ; $P(A/H_i)$ – условная вероятность события A при выполнении гипотезы H_i .

СС и СВ связаны между собой через пространство элементарных событий. При этом вероятностные свойства СВ дискретного типа описываются функцией дискретного аргумента

$$p_i = f(x_i),$$

где x_i – реализация СВ X ; p_i – соответствующая ей вероятность, а вероятностные свойства СВ непрерывного типа описываются функцией $f(x)$ (плотностью распределения), определенной на всей числовой оси, неотрицательной и нормированной на области определения.

Так, например, в биномиальном законе распределения, реализуемом в схеме независимых испытаний, пространство элементарных событий дискретной СВ X есть конечное множество целых чисел, включая 0, т. е.

$$\Omega(X) = \{0, 1, 2, \dots, n\},$$

а вероятности значений рассчитываются по формуле Бернулли:

$$P(X = n) = C_n^m p^m q^{n-m},$$

где $m = \overline{0, n}$; C_n^m – число сочетаний из n по m ; p – вероятность появления события в отдельном испытании; $q = 1 - p$.

Нормальный закон распределения – один из основных для непрерывной СВ X задается плотностью распределения:

$$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}},$$

определенной для всех и удовлетворяющей условиям $f(x) > 0$ и

$$\int_{-\infty}^{\infty} f(x) dx = 1.$$

Здесь μ_x и σ_x – соответственно математическое ожидание и среднее квадратическое отклонение СВ X . Отметим, что эти числовые характеристики являются основными для одномерной СВ X , имеющей любой закон распределения.

Универсальной формой закона распределения, имеющей место как для дискретной, так и для непрерывной СВ, является функция распределения

$$F(x) = P\{X < x\}.$$

Второй уровень теоретико-вероятностной модели, обобщающей модель одномерной СВ, связан с системой случайных величин (многомерной СВ), вероятностные свойства которой не исчерпываются свойствами отдельных величин, образующих систему, а описываются также зависимостью между ними.

Так, для двумерной непрерывной СВ (X, Y) должны рассматриваться частные плотности распределения $f_1(x)$ и $f_2(y)$, а также совместная плотность $f(x, y)$. При этом

$$f_1(x) = \int_{-\infty}^{+\infty} f(x, y) dy;$$

$$f_2(y) = \int_{-\infty}^{+\infty} f(x, y) dx.$$

Однако $f(x, y)$ не всегда может быть определена через $f_1(x)$ и $f_2(y)$. Для n -мерной СВ (X_1, X_2, \dots, X_n) функции распределения имеют гораздо более громоздкий вид, поэтому на модельном уровне при описании объекта обычно используют набор числовых характеристик.

Для двумерной случайной величины (X, Y) – это $\mu_x, \mu_y, \sigma_x, \sigma_y, r_{xy}$; для n -мерной случайной величины – это $\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n}, \sigma_{x_1}, \sigma_{x_2}, \dots, \sigma_{x_n}$, а также

корреляционная матрица $\|r_{ij}\|$, где r_{ij} – коэффициент корреляции между случайными величинами X_i и X_j .

Обобщением модели n -мерной СВ служит модель третьего уровня – случайная функция (СФ) $X(t)$, где t – вещественный параметр.

Графически СФ представима в виде набора реализаций $x_i(t)$, где каждая $x_i(t)$ – неслучайная функция ($i = \overline{1, n}$) (рис. 2.4). При фиксированном $t = t_0$ имеем сечение СФ $X(t_0)$, представляющее собой одномерную СВ.

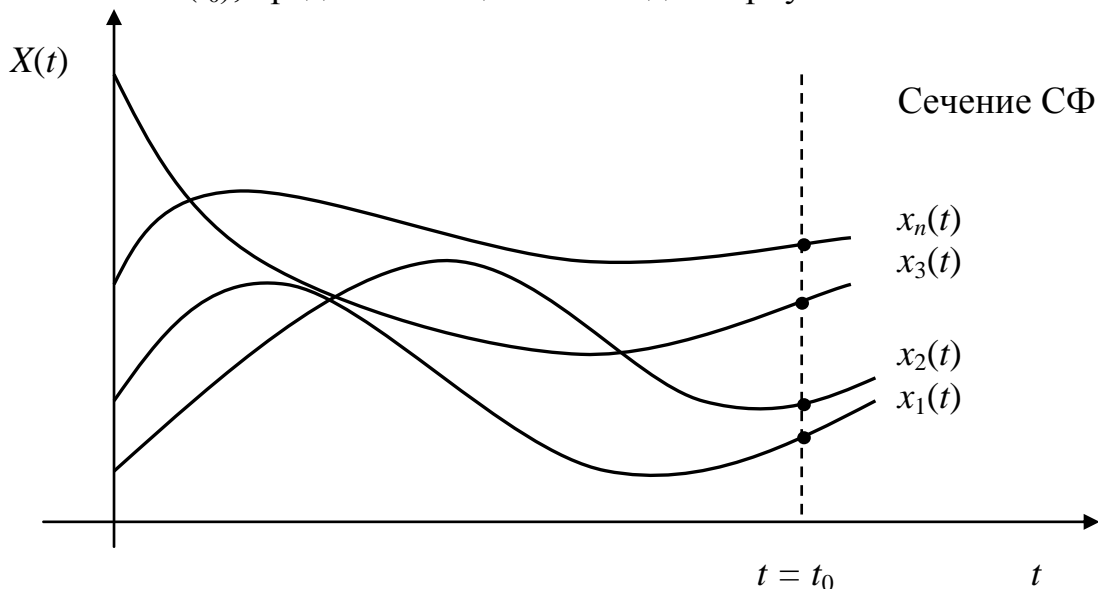


Рис. 2.4. Случайная функция $X(t)$ в виде набора реализации $x_i(t)$

Таким образом, полное вероятностное описание СФ связано с заданием бесконечномерного закона распределения всех ее сечений.

Обычно на модельном уровне ограничиваются рассмотрением СФ в рамках корреляционной теории, т. е. рассмотрением ее математического ожидания $\mu_x(t)$ и корреляционной функции $K_x(t_1, t_2)$, являющихся неслучайными функциями соответственно одного и двух аргументов.

Важным частным случаем модели СФ является модель стационарной СФ, для которой

$$\begin{cases} \mu_x(t) = \text{const}; \\ K_x(t_1, t_2) = K_x(\tau), \end{cases}$$

где $\tau = t_1 - t_2$.

Каждому из уровней теоретико-вероятностных моделей соответствует своя статистическая модель.

Поскольку математическая статистика занимается обработкой информации при наличии неопределенности, то в основу построения статистической модели должны быть положены некоторые допущения. В зависимости от их математической сути различают три основные модели математической статистики:

- 1) модель выборки;
- 2) модель регрессии;
- 3) модель случайного процесса.

Каждая из них связана с определенным исходным материалом и решает свои специфические задачи. Однако все они имеют дело с большим объемом информации и достаточно трудоемкими методами ее обработки. Поэтому реализация статистических моделей осуществляется, как правило, на компьютере, для которых разработаны стандартные программы определения статистических законов распределения, вычисления статистических характеристик СВ, проверки гипотез по критериям согласия, расчета коэффициентов регрессии, статистических характеристик случайных процессов и т. д.

В некоторых ситуациях получение исходных статистических данных путем специально организованных экспериментов невозможно. В этом случае необходимый статистический материал может быть получен с помощью специально созданных математических моделей. Их основу составляет статистическое моделирование на компьютере СС и СВ. Такой метод моделирования называется *методом статистических испытаний* (метод Монте-Карло).

Простейшей статистической моделью является одномерная *модель выборки*, в которой предполагается, что исходный статистический материал есть реализация одной СВ X с законом распределения $F(x)$.

Основой для построения модели служит простая случайная выборка, представленная в виде ряда наблюдений (табл. 2.4).

Таблица 2.4

i	1	2	...	n
x_i	x_1	x_2	...	x_n

Например, при измерении расстояния до цели каким-либо прибором за счет чисто случайных причин результат изменяется от опыта к опыту.

Реализация модели связана с построением эмпирических (статистических) законов распределения: $f^*(x)$ – статистическая плотность распределения; $F^*(x)$ – статистическая функция распределения, аналогичная теоретическим законам для СВ X . Графическим представлением $f^*(x)$ является гистограмма, а $F^*(x)$ – кривая накопленных частот.

Найденные законы содержат элемент случайности, так как они определены по конечному числу наблюдений. Поэтому для уточнения модели следует провести сглаживание статистического ряда, т. е. выбрать теоретический закон распределения, наилучшим образом описывающий исходный материал. Для этого используются критерии согласия Пирсона, Колмогорова и др.

В некоторых ситуациях ограничиваются получением точечных и интервальных оценок основных числовых характеристик μ_x и σ_x^2 .

Если результаты наблюдений зависят от некоторого параметра и изменяются от измерения к измерению не только за счет случайных причин, но и за счет существенных, то модель выборки неприменима к данному ряду наблюдений. В этом случае используют *модель регрессии*, которая предполагает, что исходный статистический материал представляет собой

реализации СВ, изменяющейся в зависимости от какого-либо параметра (времени, пространственной координаты и т. д.).

Исходный ряд наблюдений имеет вид, приведенный в табл. 2.5, где t_n – значение параметра ($i = 1, \dots, n$); x_n – значение СВ, соответствующее данному значению параметра. Например, изменение температуры воздуха в данной точке в данный момент времени зависит от высоты.

Таблица 2.5

t_i	t_1	t_2	...	t_n
x_i	x_1	x_2	...	x_n

Основным предположением модели регрессии, сводящим ее к рассмотренной ранее модели выборки, является следующее: реализации СВ могут быть представлены в виде суммы

$$x_i = f(t_i) + \delta_i,$$

где $f(t_i)$ – неслучайная функция аргумента t ; δ_i – ошибки, содержащиеся в выборке.

Другими словами, предполагается, что ряд наблюдений $\{x_1, x_2, \dots, x_n\}$ является реализацией системы независимых СВ (X_1, X_2, \dots, X_n), частные распределения которых одинаковы, за исключением математического ожидания, для которого

$$\mu_{x_i} = f(t_i).$$

Функция $f(t)$ называется сглаживающей, а ее график – линией регрессии. Основной задачей модели регрессии является определение линии регрессии $f(t)$ и оценка точности результата (рис. 2.5).

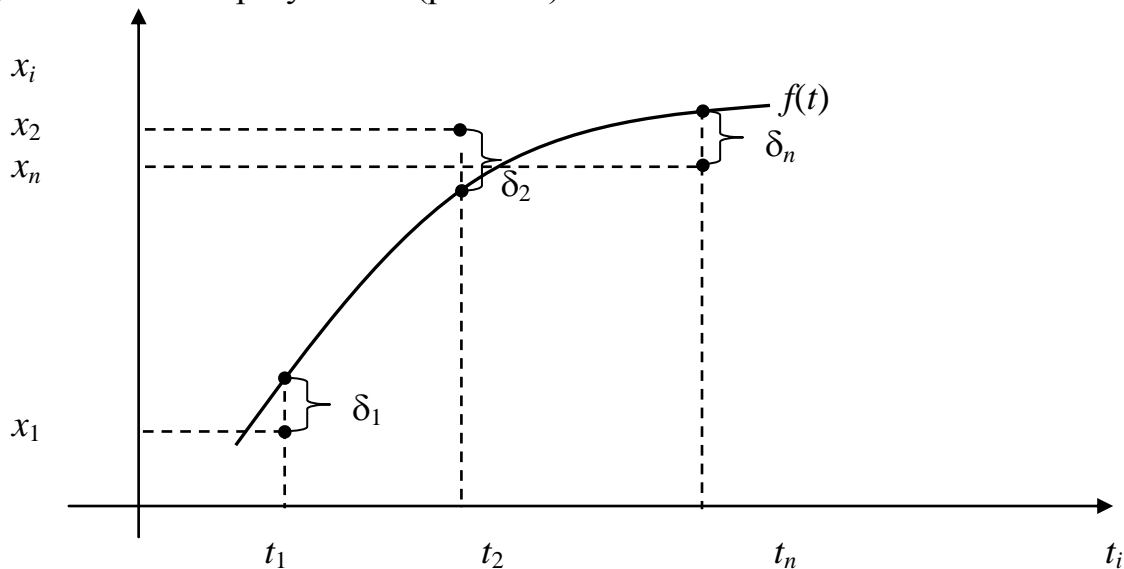


Рис. 2.5. Линия регрессии $f(t)$ и оценка точности результата δ

Сглаживающую функцию обычно задают параметрически: $f(a_1, a_2, \dots, a_m, t)$, где вид функции известен и вычислению подлежат только параметры a_1, a_2, \dots, a_m ($m < n$). Для их определения используется стандартная процедура метода наименьших квадратов (МНК), основанная на соотношении:

$$\sum_{i=1}^n [x_i - f(a_1, a_2, \dots, a_m, t)]^2 = \sum_{i=1}^n \delta_i^2 \rightarrow \min ,$$

т. е. сумма квадратов отклонений результатов наблюдений от сглаживающей кривой должна быть наименьшей.

В большинстве прикладных задач ограничиваются линейной регрессией.

Модель случайного процесса является обобщением как модели выборки, так и модели регрессии. Исходный ряд наблюдений зависит от параметра так же, как и в модели регрессии, но СВ (X_1, X_2, \dots, X_n) , реализацией которых является данный ряд, зависимы и также зависимы случайные ошибки δ_i в случайных измерениях.

Основное предположение модели случайного процесса состоит в том, что ряд наблюдений, зависящих от параметра, есть значения функции $x(t)$, являющейся реализацией СФ $X(t)$, распределение и характеристики которой заранее неизвестны.

Пусть имеется m рядов наблюдений, проведенных в одних и тех же условиях. Их можно свести в таблицу, в которой в строках расположены реализации, а в столбцах – сечения по параметру t (табл. 2.6). Данными этой таблицы могут быть, например, многолетние метеорологические наблюдения за температурой воздуха в каком-нибудь пункте в определенный месяц года.

Таблица 2.6

$x(t)$	t			
	t_1	t_2	...	t_n
$x_1(t)$	$x_1(t_1)$	$x_1(t_2)$...	$x_1(t_n)$
$x_2(t)$	$x_2(t_1)$	$x_2(t_2)$...	$x_2(t_n)$
...
$x_m(t)$	$x_m(t_1)$	$x_m(t_2)$...	$x_m(t_n)$

Обработка построенной таблицы может вестись как по «столбцам» (модель выборки), так и по «строкам» (модель регрессии). Полученные при различных значениях аргумента t результаты числовых характеристик в дальнейшем аппроксимируются по МНК аналитическими выражениями.

Если по каким-либо причинам получение нескольких статистических рядов невозможно, то для применения модели случайного процесса требуются дополнительные предположения о его характере (например, стационарность процесса).

При *статистическом* моделировании применяется метод статистических испытаний – *метод Монте-Карло*. При этом результаты, полученные при воспроизведении на имитационной модели процесса функционирования системы, являются реализациями случайных величин и функций, и для нахождения характеристик процесса требуется его многократное воспроизведение с последующей статистической обработкой информации.

Основная идея *метода Монте-Карло* заключается в следующем. Пусть необходимо определить значение детерминированной величины y . Допустим, ее можно представить в виде детерминированной функции

$$y = f(x_1, x_2, \dots, x_N),$$

где x_i – детерминированная переменная.

Заменим переменные x_1, x_2, \dots, x_N на случайные функции с известным законом распределения. Тогда значение y заменяется на случайную величину Y . Для решения задач такого вида применяется следующий алгоритм:

1) по каждой из величин x_1, x_2, \dots, x_N производится случайное испытание, в результате которого определяется некоторое конкретное значение случайной величины $x_{1i}, x_{2i}, \dots, x_{Ni}$;

2) используя найденные величины, определяется одно частное значение y_i по вышеприведенной зависимости;

3) предыдущие операции повторяются N раз, в результате чего определяется N значений случайной величины Y ;

4) на основании N значений величины Y находится ее эмпирическая функция распределения, математическое ожидание μ_Y и дисперсия σ_Y^2 .

При обработке результатов моделирования приходится довольствоваться лишь некоторыми оценками моментов, полученными на конечном числе реализации N . При независимых наблюдениях значений случайной величины Y в качестве таких оценок используются

$$\bar{Y} = \tilde{\mu}_Y = \frac{1}{N} \sum_{i=1}^N y_i,$$

$$S_b^2 = \tilde{\sigma}_Y^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{Y})^2,$$

где \bar{Y} и S_b^2 – выборочное среднее и выборочная дисперсия соответственно. Знак \sim над $\tilde{\mu}_Y$ и $\tilde{\sigma}_Y^2$ означает, что эти выборочные моменты используются в качестве оценок математического ожидания μ_Y и дисперсии σ_Y^2 .

3. ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ

Имитационная модель реального процесса (объекта, явления) – программа для компьютера, реализующая упрощенную модель этого процесса вместе с алгоритмом, описывающим течение этого процесса [10]. Когда компьютер выполняет эту программу, он «имитирует» течение реального процесса. Меняя различные параметры программы, можно имитировать течение реального процесса в различных условиях. Таким образом, возникает возможность осуществления следующего диалога:

Вопрос: (его задает исследователь, лицо, принимающее решение (ЛПР), и т. п.): «Что произойдет с объектом (процессом, явлением), если...?».

Ответ: (его дает компьютер, «проигрывая» заложенную в программу имитационную модель): «В заданных условиях с объектом (процессом, явлением) произойдет следующее...».

Организованный таким образом диалог человека с компьютером позволяет проводить те или иные эксперименты (они называются имитационными, вычислительными или машинными), получая при этом информацию, которая

может быть отнесена к реальному процессу (объекту, явлению). Разумеется, полученная информация будет иметь тем большее отношение к реальности, чем более удачно построена имитационная модель. Схема имитационных исследований представлена на рис. 3.1.

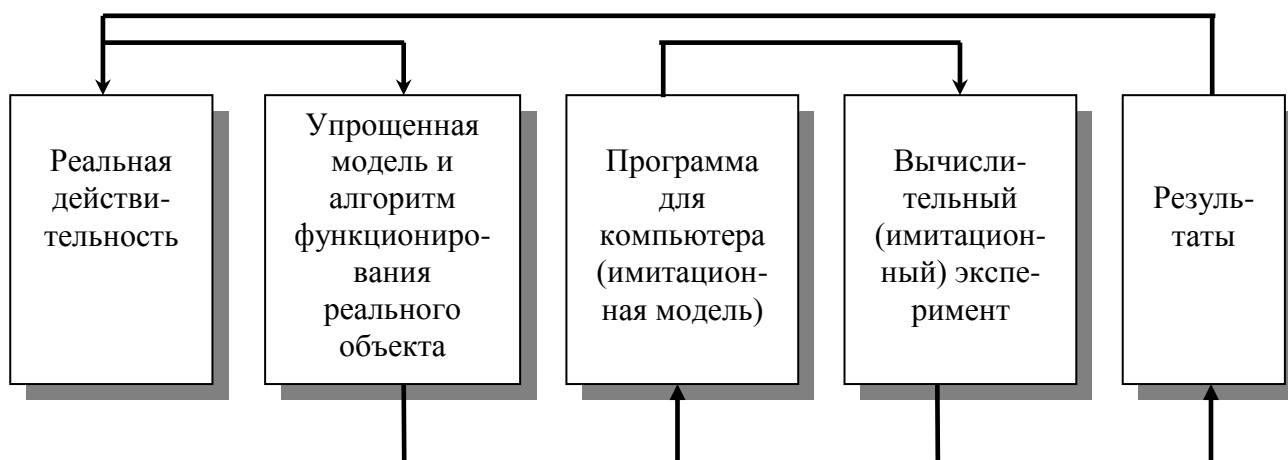


Рис. 3.1. Схема имитационных исследований

Термин «имитационное моделирование» возник как перевод английского выражения *simulation modeling*. Если обратиться к русским эквивалентам этих слов, то станет ясно, что перевод был сделан не совсем удачно, так как в нем явно имеется тавтология. Однако, поскольку термин уже получил широкое распространение, менять его едва ли возможно и целесообразно. Важнее хорошо понимать, что за ним стоит. Когда говорят «имитационная модель», обычно имеют в виду, что ей свойственны большая, чем в обычных математических моделях, близость к реальному объекту, возможность воспроизводить широкий спектр свойств объекта, использование эмпирического материала, кроме того, всегда сохраняется возможность дальнейшего уточнения модели и др. Говоря об имитационном моделировании, в него включают как процесс создания модели, так и ее исследование (проведение вычислительных экспериментов) с помощью компьютера.

Р. Шеннон [11] дает следующее определение: «...имитационное моделирование есть процесс конструирования реальной системы и постановки экспериментов на этой модели с целью оценить (в рамках ограничений, накладываемых некоторым критерием или совокупностью критериев) различные стратегии, обеспечивающие функционирование данной системы».

Как видно из этого определения, имитационное моделирование не очень приспособлено для выяснения причин того или иного явления (оно не дает ответа на вопрос «Почему?»), его роль гораздо более прагматична – давать ответ на вопрос «Что будет, если...?».

3.1. ОСНОВНЫЕ ЭТАПЫ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ

Процесс имитационного моделирования (создание модели и алгоритма функционирования реального объекта, создание программы и проведение вычислительных экспериментов) является итеративным. При этом системный

анализ чередуется и сочетается с составлением модели и алгоритма, вычислительными экспериментами, корректировкой модели и т. п.

Имитационное моделирование состоит из следующих взаимосвязанных и зачастую пересекающихся во времени этапов [10]:

- 1) постановка проблемы, формулировка цели моделирования;
- 2) системный анализ моделируемого объекта; построение концептуальной модели;
- 3) составление модели и алгоритма функционирования реального объекта, при необходимости – структуризация модели;
- 4) программная реализация и вычислительные эксперименты;
- 5) анализ результатов экспериментов и коррекция модели, алгоритма, программы;
- 6) проведение имитационных экспериментов в целях решения поставленной проблемы.

Охарактеризуем каждый этап.

1. Имитационное моделирование, как и всякое другое, должно начинаться с постановки проблемы, с ясного и четкого понимания целей моделирования. Желательно перечислить те вопросы, ответы на которые должны быть получены в ходе вычислительных экспериментов.

Необходимость построения имитационной модели, как правило, возникает в процессе прикладных исследований конкретного объекта. Поэтому одной из важнейших особенностей этого этапа моделирования является непременно участие в разработках человека (или группы людей), который будет использовать будущую модель для решения поставленной проблемы, т. е. ЛПР. Нужно отметить, что в реальной жизни цели редко бывают четко очерченными. Это существенно затрудняет точную постановку проблемы. Поэтому упомянутый выше список вопросов следует обязательно согласовывать с ЛПР. С ним же следует решать вопрос о масштабах задачи: о ее объеме, границах в пространстве, продолжительности моделируемого отрезка времени (интересуют ЛПР долгосрочные или краткосрочные эффекты) и т. д.

Необходимо уточнить объект исследования: систем в целом или отдельные ее части. Лишь после этого целесообразно переходить к последующим этапам.

2. С учетом целей моделирования следует выявить существенные особенности изучаемого объекта, осуществить отбор необходимых сведений о нем и построить так называемую концептуальную модель объекта. Эта модель представляет формализованное описание объекта (словесное описание, представление в виде схем, диаграмм и т. п.), являющееся основой для создания имитационной модели.

Построение концептуальной модели предполагает:

- а) выявление основных процессов, которые должны быть учтены при моделировании;
- б) выявление тех основных характеристик объекта, которые необходимо иметь для решения исходной проблемы;
- в) определение множества переменных и параметров, которые влияют на динамику этих характеристик;

- г) определение множества входных и выходных данных модели;
- д) установление границ и законов взаимодействия объекта с окружающей средой (в частности, определение законов случайных воздействий на объект);
- е) разработку причинно-следственных связей, временных отношений и гипотез, согласно которым осуществляется взаимное объединение всех перечисленных компонент в единую систему – имитационную модель.

3. Далее осуществляется переход от качественных зависимостей концептуальной модели к точному алгоритмическому описанию. Исходным пунктом при этом является задание вектора состояния модели, компоненты которого – это те характеристики изучаемого объекта, которые выделены при построении концептуальной модели как базовые, несущие в себе необходимую информацию для решения поставленной проблемы

$$Z^t = (Z_1^t, Z_2^t, \dots, Z_n^t).$$

Часто бывает полезным сразу выделять «управляемые компоненты» вектора состояния. В общем случае каждая компонента вектора Z^t есть функция F , зависящая от времени, а также от множества значений параметра $\{H\}$, некоторого подмножества компонент вектора состояния Z^t , множества внешних факторов V^t и управляющих воздействий X^t , т. е.

$$Z^t = F(X^t, V^t, H, Z^t, t).$$

В зависимости от цели моделирования (целевой функции), от проблемы, которую необходимо разрешить с помощью модели, вводится так называемое системное время, моделирующее ход времени в реальной системе. Различают два типа шкал модельного времени: равномерный и событийный. Для первого типа характерно введение некоторого постоянного шага Δt изменения времени. В этом случае вектор состояния модели рассматривается в моменты $t + k\Delta t$, где k – целое положительное число (разумеется, величина $t + k\Delta t$ не должна превосходить T – величину моделируемого промежутка времени). В этом случае каждому моменту реального времени t_i^p ставится в соответствие момент модельного времени t_i^m , причем

$$t_1^m - t_2^m = A(t_1^p - t_2^p),$$

где A – масштабный коэффициент пропорциональности.

Если моделируемый объект изменяется лишь при наступлении некоторого события, а в остальные моменты времени остается без изменений, пользоваться равномерной шкалой неудобно, поскольку вектор состояния на отрезке времени между двумя событиями остается постоянным. В этом случае отсчет времени в модели ведется «по событиям», т. е. каждый последующий момент времени в модели наступает только тогда, когда в ней моделируется наступление некоторого события.

После, а иногда и параллельно с заданием вектора состояния системы и выбором временного шага производится декомпозиция модели и выявление ее блочной конструкции (если в этом есть необходимость). С этой целью множество переменных модели делится на непересекающиеся подмножества, в каждое из которых входит группа «однородных» переменных. Понятие «однородности», как правило, определяется самим исследователем в

зависимости от цели моделирования, количества и качества исходных данных. Чаще всего однородными, входящими в один блок считаются переменные, описывающие отдельный процесс, какую-либо подсистему или элемент исходного объекта, группу факторов, имеющих одну и ту же природу (например, климатообусловленные или антропогенные факторы и т. п.). В результате декомпозиции модель представляется в виде комплекса взаимосвязанных подмоделей – блоков, которые взаимодействуют по определенным законам и в итоге позволяют провести имитационное исследование объекта.

Блочный принцип построения модели имеет целый ряд преимуществ, особенно ощутимых при создании сложных имитационных моделей. Прежде всего, если модель достаточно сложна, требует значительного объема памяти и машинного времени, то возможности использования такой модели для имитационного эксперимента оказываются весьма ограниченными. Единственным выходом в этом случае оказывается декомпозиция модели для того, чтобы программы, реализующие отдельные блоки, работали последовательно и обменивались информацией по тем или иным правилам.

Поскольку блоки описывают различные подсистемы, процессы, факторы, системное время для каждого из них может быть различным. Оно определяется, исходя из внутренних потребностей блока. Если, к примеру, в блоке описываются «быстрые» изменения, то время исчисляется мелкими единицами измерения: секундами, часами, сутками (скажем, при моделировании процесса фильтрации воды в почве после дождя, процесса накопления автомобилей перед светофором и т. п.); если изменения объекта во времени происходят достаточно медленно (рост дерева, старение оборудования, глобальные изменения климата земли и т. п.), то время исчисляется месяцами, годами, десятилетиями и даже тысячелетиями. Однако необходимо позаботиться о том, чтобы в итоге все результаты по отдельным блокам были сведены к единому системному времени, принятому для модели в целом.

После завершения декомпозиции модели следует приступить к разработке отдельных ее блоков. Для каждого из них:

а) уточняются и конкретизируются те гипотезы, которые непосредственно относятся к процессам, аспектам, элементам, «принадлежащим» данному блоку;

б) определяется соответствующее подмножество входных и выходных данных, причем эти данные могут принадлежать как к множеству «входов» и «выходов» общей модели, так и к множеству локальных (или внутренних) входных и выходных данных других блоков;

в) формируется множество параметров;

г) формализуются основные законы взаимодействия элементов блока.

При этом происходит переход от качественных зависимостей концептуальной модели к точным количественным зависимостям и логическим схемам взаимодействия элементов внутри блока.

Что касается последнего, то в практике имитационного моделирования далеко не все связи концептуальной модели удастся отразить в виде

зависимостей теоретического характера. Часто приходится вводить эмпирические зависимости, полученные на основе данных натурных наблюдений, в результате обобщения опыта моделирования подобных объектов и т. п. Неоценимую роль на этапе формализации играет экспертная оценка полученных эмпирических зависимостей, коэффициентов и т. п.

Существенное влияние на качество разработки отдельных блоков оказывает выбор различных математических средств моделирования: аппарата дифференциальных уравнений, статистического моделирования, методов теории оптимального управления, алгоритмизации логических конструкций и т. п. Важным также является то, в каком сочетании эти средства используются для решения исходной проблемы, как согласуется выбор методов с возможностями компьютера. Поэтому на этапе разработки блоков необходимо участие не только математиков и экспертов, но и системного программиста, знакомого с характеристиками вычислительной техники, математическим и техническим обеспечением процесса счета. Это позволит оценить возможности программы, реализующей на компьютере создаваемую модель, определить пути возможно нежелательных, но часто так необходимых упрощений.

4. Заключительным в построении модели является объединение блоков в имитационную модель на базе стандартного или специально созданного математического обеспечения. Здесь особую роль играет выбор языка программирования: либо это будут универсальные языки типа Си, Паскаль, Фортран и т. п., либо специализированные языки имитационного моделирования типа ДИНАМО, GPSS, Симула и т. д. Проблема носит принципиальный характер, так как специализированные языки удобны для программирования, отличаются концептуальной направленностью, что позволяет при составлении несложных имитационных моделей опустить этап построения формальной модели. Однако такие языки требуют специальных трансляторов, которые не всегда входят в стандартное математическое обеспечение компьютера. В свою очередь, использование универсальных языков, как правило, сильно увеличивает объем программ, делает их громоздкими и «трудночитаемыми» для пользователя, хотя и допускает проведение имитационных экспериментов практически на любых компьютерах, что расширяет область практического применения модели.

Составлению машинных программ, реализующих всю модель в целом, предшествуют испытания и отработка различных схем взаимодействия блоков. Здесь удобно бывает рассматривать имитационную модель как коллектив автоматов с памятью и без нее, детерминированных или стохастических, а работу модели – как изучение с помощью компьютера коллективного поведения автоматов в случайной или детерминированной среде.

На этом заканчиваются первые четыре этапа построения имитационной модели. Они представлены на рис. 3.2.

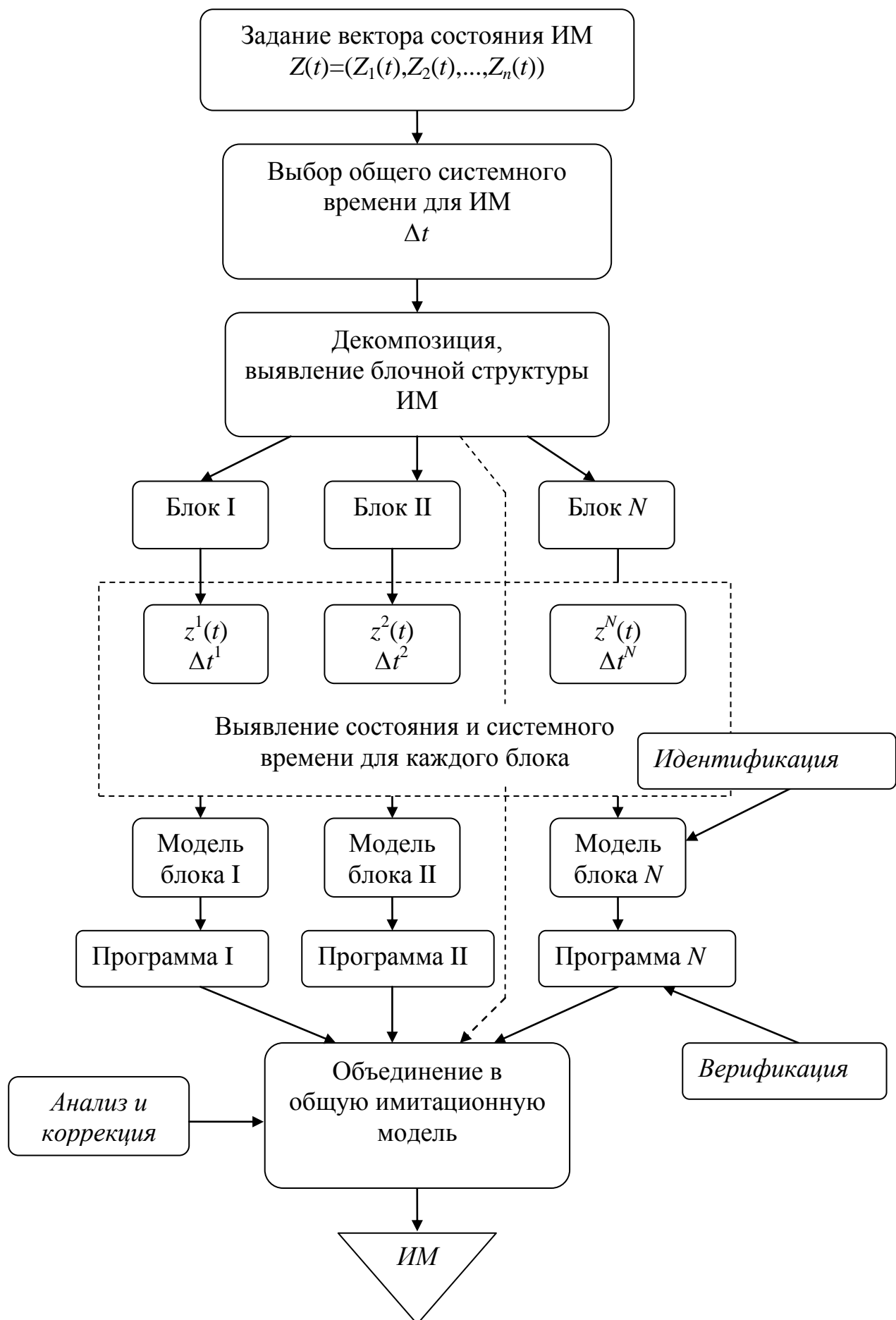


Рис. 3.2. Первые четыре этапа построения имитационной модели

5. Маловероятно, что построенная имитационная модель сразу же окажется удачной. Скорее всего, в ней обнаружатся ошибки. Для отыскивания этих ошибок и используются вычислительные эксперименты как с отдельными блоками, так и с системой в целом. Причем желательно проведение таких экспериментов, результат которых по тем или иным причинам может быть предсказан. Тогда отклонение результата от прогнозируемого служит индикатором наличия ошибок либо в модели, либо в алгоритме, либо в программе, либо сразу в нескольких местах.

Устранение ошибок продолжается до тех пор, пока результаты экспериментов придут в соответствие с прогнозируемыми. После этого можно считать, что имитационная модель создана.

Однако имитационное моделирование состоит не только в построении самой имитационной модели, т. е. инструмента для проведения научных исследований, но и в применении этого инструмента. Необходимо уметь использовать эту модель в исследовательских целях – тех целях, для которых она создавалась.

3.2. ИДЕНТИФИКАЦИЯ И ВЕРИФИКАЦИЯ ИМИТАЦИОННОЙ МОДЕЛИ

Итак, имитационная модель построена. Однако прежде чем ее использовать, необходимо решить следующие задачи:

1. Выбрать числовые значения неопределенных пока числовых параметров (идентификация).

2. Убедиться, что при этих значениях параметров модель хорошо соответствует моделируемому объекту, адекватна ему (верификация, проверка адекватности).

Приведем наиболее распространенную постановку задачи идентификации.

Пусть $\vec{y}(t) = \vec{y}(t, \vec{h})$ – вектор выходных характеристик имитационной модели, зависящей от набора параметров $\vec{h} = \{h_1, h_2, \dots, h_m\}$. Предположим, что имеется ряд известных векторов $\vec{y}(t)$, $t \in [t_1, t_2]$, представляющих собой результаты натурных наблюдений за моделируемым объектом. Разобьем промежуток $[t_1, t_2]$ на две части: $[t_1, \tau]$ – обучающий промежуток и $[\tau, t_2]$ – экзаменующий промежуток.

Задача идентификации состоит в том, чтобы найти такой набор параметров \vec{h} , который доставляет минимум функционалу

$$\sum_{t=t_1}^{\tau} (\vec{y}(t) - \vec{y}(t, \vec{h}))^2.$$

Для решения этой задачи могут быть использованы известные численные методы нахождения экстремума функции многих переменных.

Обозначим найденный в результате решения этой задачи наилучший набор параметров через \vec{h}^* . Используя в модели именно эти параметры, получим

выходные характеристики в виде $\vec{y} = \vec{y}(t, \vec{h}^*)$. На этом решение задачи идентификации закончено.

Перейдем к задаче верификации, которая мало формализована. Прежде всего убедимся, что на экзаменуящем промежутке времени $[\tau, t_2]$ расчетная траектория $\vec{y} = \vec{y}(t, \vec{h}^*)$ близка к фактической.

Сравнение этих двух траекторий позволяет судить об адекватности модели. Существуют специальные методы оценки близости траекторий. Один из них состоит в вычислении коэффициента несовпадения

$$U = \frac{\sqrt{\frac{1}{T} \sum_{t=1}^T (\vec{y}(t) - \vec{y}^*(t))^2}}{\sqrt{\frac{1}{T} \sum_{t=1}^T \vec{y}^2(t) + \frac{1}{T} \sum_{t=1}^T (\vec{y}^*)^2(t)}}, 0 \leq U \leq 1.$$

Чем ближе U к нулю, тем ближе модельная траектория к фактической. В случае, когда U равен единице, модель не является адекватной и требует либо перестройки структуры, замены или уточнения гипотез, либо идентификации по более полным и достоверным данным.

Однако даже если близость траекторий имеет место, еще нет гарантии того, что модель адекватна реальному объекту. Естественно потребовать, чтобы выполнялись следующие условия.

1. Машинная реализация соответствует формальной модели.
2. «Динамика» модели соответствует «динамике» реального объекта.
3. Результаты моделирования правильно интерпретируются.

Проверка адекватности в этом смысле осуществляется на основе экспертного анализа и статистических методов.

Под соответствием машинной и формальной моделей понимается, во-первых, идентичность их алгоритмических структур (сохраняется ли логика построения модели при машинной реализации?) и, во-вторых, совпадение областей варьирования компонент вектора состояния формальной и машинной моделей. Второе требование означает, что численные методы, которые используются для реализации модели на компьютере, не должны давать такую погрешность, которая выводит некоторую компоненту из области допустимых значений. Например, если шаг интегрирования системы обыкновенных дифференциальных уравнений с положительно-определенными переменными выбран слишком большим, то можно получить отрицательные значения искомых переменных, что не соответствует их смыслу.

Соответствие имитируемой и реальной динамики частично проверяется уже на стадии верификации модели. Однако известно, что практически любую имитационную модель можно отладить до нужной степени совпадения моделируемой и фактической траекторий. На рис. 3.3 приведен пример прогнозирования по модели, которая хорошо аппроксимирует фактическую траекторию на ретроспективном периоде, но дает абсолютно неверный прогноз, т. е. неадекватно отражает поведение реального объекта.

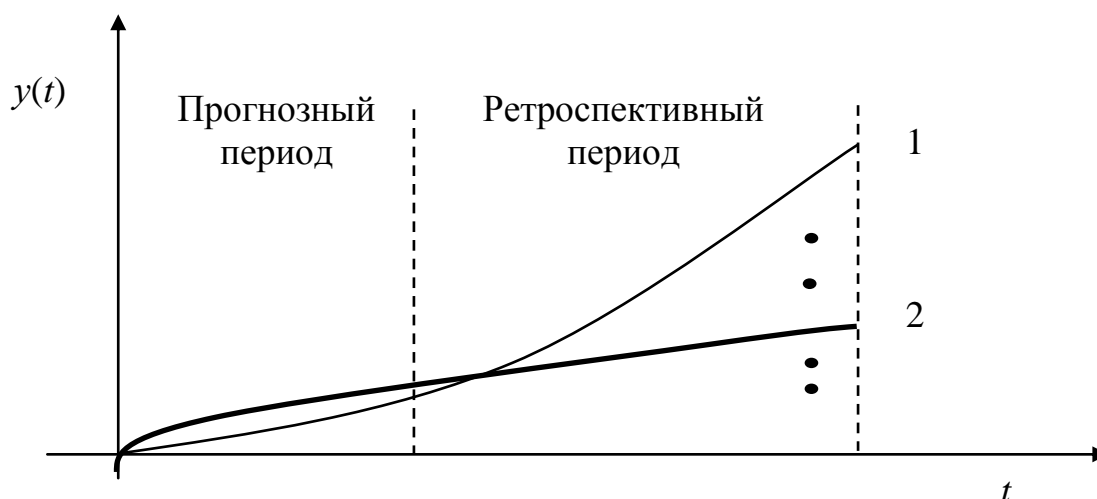


Рис. 3.3. Прогнозирование по модели:
1 – фактическая траектория; 2 – расчетная траектория

Следовательно, для того чтобы оценить адекватность модели, необходимо качественно исследовать ее динамические свойства путем проведения серии *тестовых расчетов*. Можно выделить два основных типа тестов. К первому относятся тесты, в которых задаются правдоподобные значения входов и управляющих воздействий. Если в этом случае расчеты по модели не противоречат известным законам поведения реального объекта (нашим представлениям о нем), то это говорит в пользу адекватности модели. В противном случае необходимо найти причины несогласованности и перестроить модель. Второй тип тестов основан на использовании критических ситуаций, т. е. данных, которые не характерны для исследуемого объекта, но тем не менее могут иметь место. Эти тесты особенно важны для сложных моделей, предназначенных для долгосрочного прогнозирования, поскольку чем лучше модель описывает поведение объекта в критических условиях, тем больше можно быть уверенным в правильности расчетов для нормальных условий.

Адекватность модели характеризуется также ее *чувствительностью* по отношению к изменениям параметров и начальных значений вектора состояния. Если результаты исследований с помощью модели существенно изменяются при малых возмущениях параметров и незначительных отклонениях от начальных данных, другими словами, если модель не является устойчивой, то ее нельзя считать адекватной (при условии, конечно, что исследуемый объект обладает устойчивостью в этом смысле). Требование устойчивости тем более важно, чем менее точно могут быть определены параметры модели.

Анализируя адекватность модели, следует помнить о цели исследования, о проблеме, для решения которой разрабатывается модель. Дело в том, что конкретная модель может вполне отвечать одной цели и быть совершенно непригодной для решения других задач. Здесь важен вопрос об области применимости модели. Модель *пригодна* только тогда, когда она дает возможность реализовать цели исследования.

Основными методами оценки адекватности и пригодности модели являются, безусловно, неформальные (метод экспертного оценивания, соблюдение принципа «здравого смысла» и т. п.). Однако в некоторых случаях удается формализовать эту процедуру. В частности, весьма полезными оказываются различные статистические методы, спектральный анализ и др.

Итак, модель идентифицирована, верифицирована, проверена на адекватность. Если все этапы выполнены успешно, то она является готовым инструментом исследования поставленной проблемы и можно переходить к главному этапу имитационного моделирования – проведению имитационного эксперимента, которое сопровождается, с одной стороны, планированием эксперимента, а с другой – обработкой результатов этого эксперимента.

Эксперименты делятся по цели исследования на два основных типа: дескрипторные и оптимизационные. Эксперименты первого типа проводятся в целях исследования объекта. Другой тип включает эксперименты, направленные на выявление наилучших стратегий управления исходным объектом.

Результаты имитационных экспериментов должны быть обработаны специальными методами и представлены пользователю в удобном для него виде.

3.3. ПЛАНИРОВАНИЕ ИМИТАЦИОННЫХ ЭКСПЕРИМЕНТОВ

Проведение имитационного эксперимента – один из главных этапов имитационного исследования, включающий планирование эксперимента, расчеты по модели и обработку результатов эксперимента. Сразу же возникает вопрос о том, при каких внешних воздействиях проводить расчеты, сколько расчетов проводить для того, чтобы быть уверенным в достоверности полученного решения и т. д. Все эти проблемы могут быть решены в процессе планирования эксперимента.

Цель эксперимента – установить связь между воздействиями на модель и ее откликом на это воздействие. На этом этапе модель можно представить в виде

$$y = f(x),$$

где x – воздействие на модель или фактор; y – результат воздействия или реакция; f – поверхность реакции.

В общем случае x и y есть вектор-функции, зависящие от времени. Любой имитационный эксперимент в этом случае может быть направлен либо на исследование поверхности реакции (*задачи прогнозирования*, например), либо на поиск максимума или минимума поверхности реакции в некотором пространстве факторов (*задачи оптимального управления объектом* и т. п.).

Факторы могут быть либо количественными, либо качественными. Рассмотрим только факторы первого типа.

Пусть для простоты поверхность реакции описывается функцией, зависящей от двух количественных факторов:

$$f = f(x_1, x_2),$$

причем x_1 и x_2 могут принимать дискретные фиксированные значения из областей X^1 и X^2 . Значения фактора назовем уровнем, а совокупность всех возможных пар (x_1, x_2) – полным факторным планом (рис. 3.4).

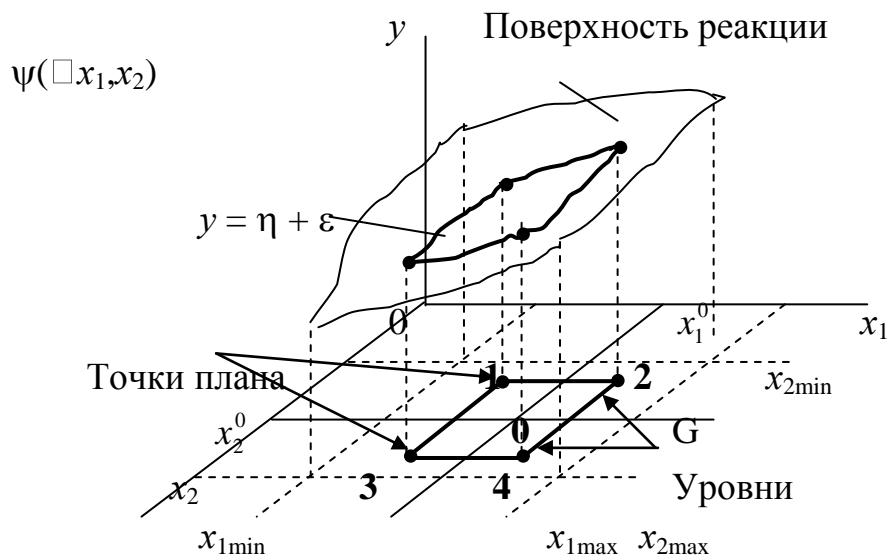


Рис. 3.4. Полный факторный план

Чем больше точек плана будет рассмотрено, т. е. чем полнее построен план, тем точнее представления о виде поверхности реакции. Однако, несмотря на это, применение полных факторных планов ограничено и возможно лишь в случае незначительного числа факторов и их уровней. Более часто используются неполные факторные планы, требующие меньшего числа точек плана и не приводящие при этом к ощутимым потерям информации о поверхности реакции. Здесь в основном исследуются несколько главных факторов, а неполные факторные планы применяются для «отсеивания» несущественных факторов. Процедура отсеивания состоит в последовательном построении неполных планов.

Рассмотрим, например, двухфакторную модель

$$f = f(x_1, x_2),$$

где факторы x_1, x_2 имеют не более 2 уровней каждый. Сначала имитационный эксперимент проводится для начальных значений этих факторов (x_1^0, x_2^0) , а затем задается некоторое фиксированное изменение каждого уровня δ_1 и δ_2 , где

$$\delta_1 = \frac{x_1^0}{n}, \quad \delta_2 = \frac{x_2^0}{n}, \quad n - \text{число уровней для } x_1 \text{ и } x_2 \text{ (обычно одинаковое),}$$

и строятся неполные факторные планы. Реализация этих планов дает четыре точки поверхности реакции $f_{+\delta_1}^1, f_{-\delta_1}^1, f_{+\delta_2}^2, f_{-\delta_2}^2$, по которым можно судить о степени влияния каждого фактора. Так, если

$$|f_{+\delta_1}^1 - f_{-\delta_1}^1| > |f_{+\delta_2}^2 - f_{-\delta_2}^2|,$$

то x_1 является более существенным фактором, и наоборот.

Еще одна типичная задача планирования эксперимента состоит в аппроксимации истинной поверхности реакции некоторой функцией ϕ ,

зависящей от тех же факторов. Как правило, удается построить линейную зависимость $\varphi = \sum_{i=1}^k a_i x_i + a_0$, где a_i – коэффициенты линейного многочлена.

Пусть, например, количество факторов $k = 2$. Тогда для построения полинома

$$\varphi = a_1 x_1 + a_2 x_2 + a_0$$

требуется проведение имитационного эксперимента по полному двухфакторному плану с уровнями $x_1^0 \pm \delta_1$, $x_2^0 \pm \delta_2$.

В случае неудовлетворительной аппроксимации есть возможность строить полиномы более высокой степени. Например, полином второй степени

$$\varphi = a_0 + a_1 x_1 + a_2 x_2 + a_{11} x_1^2 + a_{12} x_1 x_2 + a_{22} x_2^2 + a_{112} x_1^2 x_2 + a_{122} x_1 x_2^2 + a_{1122} x_1^2 x_2^2.$$

Его коэффициенты находятся на основе новых экспериментов за счет дополнительных уровней. Процесс продолжается до тех пор, пока аппроксимация не даст удовлетворительных результатов.

Описанный метод называется методом поверхности реакции.

Более узкой является проблема поиска экстремумов поверхности реакции, для решения которой используются известные методы оптимизации на заданном множестве значений факторов. Среди них отметим метод наискорейшего спуска (или подъема), который состоит в исследовании поверхности реакции в окрестности некоторой точки с помощью линейных аппроксимирующих поверхностей-гиперплоскостей. Такие гиперплоскости обычно строятся с помощью простых экспериментов, как правило, однофакторных (рис. 3.5).

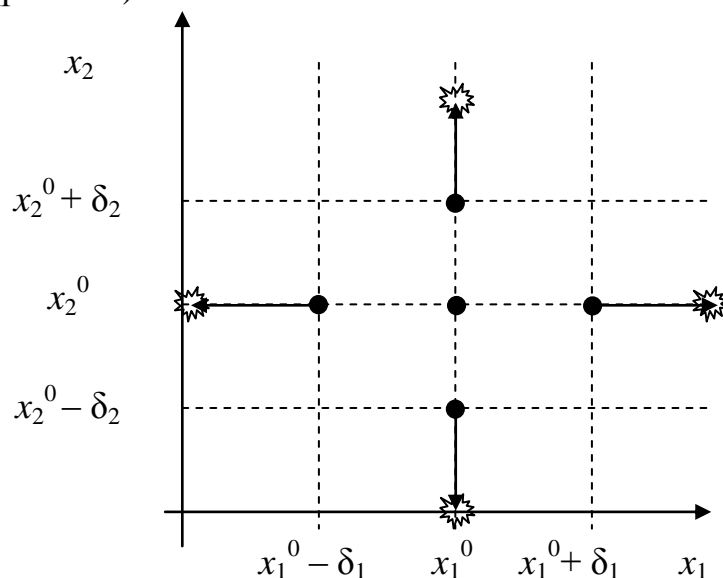


Рис. 3.5. Однофакторный эксперимент:

• – начальные изменения; * – дополнительные изменения

По построенной гиперплоскости определяется направление движения к точке оптимума, а затем в этом направлении делается небольшой «шаг». Далее процедура повторяется.

Метод наискорейшего спуска не гарантирует минимума. Если допустить, что поверхность реакции имеет несколько локальных минимумов, то целесообразно несколько раз применять этот метод, отталкиваясь всякий раз от сильно отличающихся начальных условий.

Заметим, что линейная аппроксимация вблизи точки оптимума оказывается неэффективной. В окрестности точки оптимума, где поверхность реакции почти стационарна, используется аппроксимация более высокого порядка, например квадратичным полиномом.

Все перечисленные методы планирования эксперимента относятся к детерминированным моделям. Для стохастических моделей однократная реализация построенного плана не позволяет получить желаемую информацию об изученной поверхности реакции. В этом случае необходимо несколько раз реализовывать один и тот же план с различными начальными состояниями генератора случайных чисел (каждая реализация называется репликой). Определение объема выборки (количества реплик) в имитационном эксперименте представляет собой очень трудную, но важную задачу. С одной стороны, увеличение объема ведет к увеличению затрат машинного времени и тем самым денежных средств. С другой стороны, чем больше количество реплик, тем более достоверна информация, полученная с помощью модели, и меньше возможные потери, обусловленные использованием недостоверной информации. Минимизация суммарных потерь всякий раз осуществляется с помощью методов статистического анализа и методов оптимизации.

4. КОРРЕЛЯЦИОННЫЙ АНАЛИЗ

4.1. КОРРЕЛЯЦИЯ И ПРИЧИННАЯ ЗАВИСИМОСТЬ

Независимо от числа исследуемых признаков цель корреляционного анализа – получение объективных сведений о характере взаимодействия между изучаемыми показателями необходимых для проверки существующих или выдвижения новых гипотез о внутренних механизмах интересующих нас процессов.

Ввиду того что измеряемые в экспериментах признаки являются случайными величинами в силу присущей всем естественным объектам изменчивости, изучение их совместного варьирования представляет собой задачу об определении статистической зависимости.

Возможны две постановки этой задачи [10]. Нас может интересовать либо взаимозависимость между измеряемыми величинами, либо зависимость одной или нескольких величин от остальных. Например, можно рассмотреть вопрос, существует ли связь между среднегодовой температурой и географической широтой городов на территории России. При такой постановке – это задача о взаимозависимости, но если мы хотим, используя данные о географической широте, получить информацию о среднегодовой температуре, то приходим к задаче о зависимости. Отличие возникает за счет того, что в первом случае обе используемые величины – случайные, а во втором – случайными являются

только значения среднегодовой температуры, а значения географических широт фиксированы, например с шагом в 5°.

Исследование взаимозависимости приводит к теории корреляции, исследование зависимости – к теории регрессии.

Сделаем одно замечание, имеющее принципиальное значение при изучении статистической зависимости.

Статистическая зависимость, как бы сильна она ни была, никогда не может установить наличия причинной связи. Соображения о причинно-следственных отношениях не могут и не должны быть выведены только из анализа экспериментальных данных. Для формулирования гипотез об этих соотношениях должны привлекаться сведения извне, т. е. основанные на теоретических представлениях о сущности изучаемых процессов.

Конечно, корреляционные показатели можно использовать при изучении и причинных связей, но только при экспериментальной проверке вытекающих из теоретических предпосылок следствий.

К сожалению, иногда показатели корреляции используются для неправильных утверждений о якобы имеющей место причинной зависимости, что не всегда правомерно. Изучение совместной вариации двух или более случайных величин, а именно это и составляет суть корреляционного анализа, не дает логических оснований для утверждения о существовании между этими величинами причинной связи.

4.2. КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ

Как уже было указано выше, при проведении экспериментов или наблюдений очень часто измеряются две или более характеристики объекта исследования. При этом значительный интерес представляет изучение численной меры степени взаимосвязи. Рассмотрим одну из таких мер, предназначенную для оценки взаимозависимости между двумя переменными, введенную в статистическую практику К. Пирсоном.

Пусть при исследовании технико-экономических показателей транспортной сети получены данные, приведенные ниже.

Среднемесячная температура воздуха, °С	Месячный доход на маршруте, тыс. руб.	Среднемесячная температура воздуха, °С	Месячный доход на маршруте, тыс. руб.
-22,00	369,852	15,42	538,146
-17,89	463,974	15,00	547,134
-4,10	570,15	8,40	587,776
10,33	584,164	1,32	575,716
20,13	508,522	-3,50	539,568
21,04	532,252	-15,16	634,032

Связь между среднемесячной температурой воздуха и месячным доходом на маршруте можно выразить графически в виде диаграммы рассеяния (рис. 4.1), где каждый месяц в году характеризуется точкой, координаты которой x и y равны соответственно среднемесячной температуре воздуха и месячному доходу на маршруте трамвая.

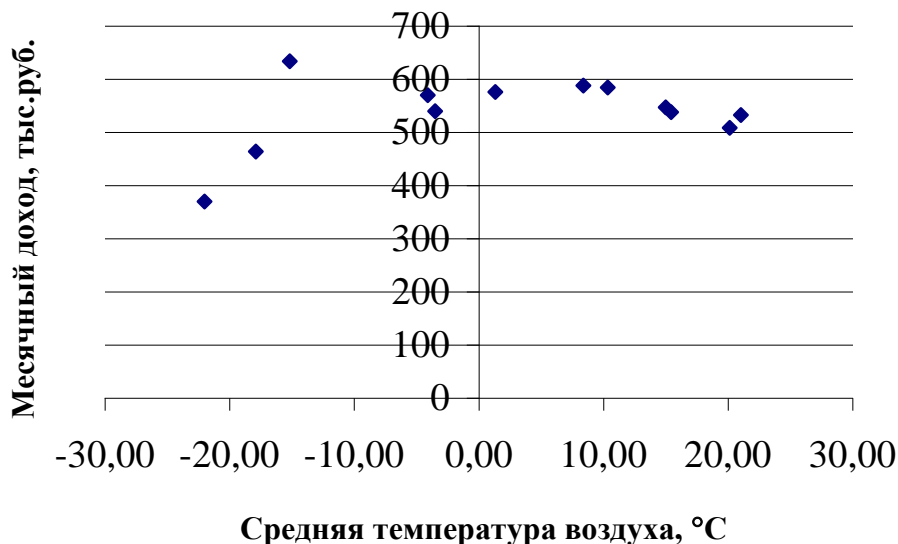


Рис. 4.1. Диаграмма рассеяния

О взаимозависимости между парой переменных можно говорить тогда, когда уменьшению (увеличению) одной из них будет соответствовать уменьшение (увеличение) другой либо когда уменьшению (увеличению) первой будет соответствовать увеличение (уменьшение) второй переменной. В первом случае будем говорить о положительной корреляции (прямая зависимость), во втором – об отрицательной (обратная зависимость).

Дадим количественную оценку этих взаимодействий. Из диаграммы рассеяния видно, что положение точек по каждой из координат характеризуется величиной и знаком отклонения от средней величины, т. е. $x_j - \bar{x}$ и $y_j - \bar{y}$. Рассмотрим произведение этих отклонений. Если объект имеет по обеим переменным значения, большие или меньшие средней величины, то произведение отклонений будет положительным, так как знаки отклонений одинаковы.

Если изменения переменных разнонаправлены, то положительным значениям $x_j - \bar{x}$ будут соответствовать в основном отрицательные значения $y_j - \bar{y}$ и наоборот.

Суммировав произведения отклонений по всем объектам, получим величину $\sum_{j=1}^n (x_j - \bar{x})(y_j - \bar{y})$, которая будет велика и положительна, когда x и y сильно связаны прямой зависимостью, и велика и отрицательна, когда x и y сильно связаны обратной зависимостью.

В тех случаях, когда между x и y нет систематической связи, т. е. большим значениям x одинаково часто соответствуют как большие, так и малые значения y , величина $\sum_{j=1}^n (x_j - \bar{x})(y_j - \bar{y})$ должна быть достаточно близкой к нулю.

Сама по себе величина $\sum_{j=1}^n (x_j - \bar{x})(y_j - \bar{y})$ непригодна в качестве меры связи, так как она зависит от числа пар значений, учитываемых при расчетах. А так как во многих случаях желательно сравнить степень связи между x и y в выборках разного объема, то показатель взаимосвязи не должен зависеть от него. Это достигается усреднением по объему выборки. По той же причине, что и при расчете дисперсии, следует разделить сумму произведений отклонений на $n - 1$. Полученная величина носит название *ковариации*

$$s_{xy} = \frac{\sum_{j=1}^n (x_j - \bar{x})(y_j - \bar{y})}{n - 1}. \quad (4.1)$$

Признаки x и y , между которыми рассчитывается взаимосвязь, могут измеряться в разных единицах, иметь произвольные средние и дисперсии. Вычитание соответствующих средних из x и y делает s_{xy} независимым от средних. Если теперь разделить ковариацию на произведение стандартных отклонений, то получится безразмерный коэффициент связи, который носит название *коэффициента корреляции*

$$r_{xy} = s_{xy} / s_x s_y, \quad (4.2)$$

где s_x , s_y – средние квадратические отклонения для x и y .

Ясно, что $s_{xy} = s_{yx}$ и $r_{xy} = r_{yx}$. С помощью коэффициента корреляции можно численно оценивать величину взаимосвязи. Несложные преобразования позволяют привести формулу (4.2) к виду, рациональному для вычислений и программирования:

$$r_{xy} = \frac{n \sum_{j=1}^n x_j y_j - (\sum_{j=1}^n x_j)(\sum_{j=1}^n y_j)}{\sqrt{(n \sum_{j=1}^n x_j^2 - (\sum_{j=1}^n x_j)^2)(n \sum_{j=1}^n y_j^2 - (\sum_{j=1}^n y_j)^2)}}. \quad (4.3)$$

Используя эту формулу, рассчитаем коэффициент корреляции между среднемесячной температурой воздуха и месячной выручкой на маршруте.

Промежуточные вычисления, с учетом нормирования температуры относительно -30°C :

$$\begin{aligned} \sum_{j=1}^n x_j &= 388,99; & \sum_{j=1}^n y_j &= 6451,286; & \sum_{j=1}^n x_j^2 &= 15091,92; \\ \sum_{j=1}^n y_j^2 &= 3519284; & \sum_{j=1}^n x_j y_j &= 212935,3. \end{aligned}$$

Окончательно имеем:

$$r_{xy} = \frac{12 \cdot 212935,3 - 388,99 \cdot 6451,286}{\sqrt{(12 \cdot 15091,92 - (383,99)^2)(12 \cdot 3519284 - (6451,286)^2)}} = 0,3387.$$

Таким образом, оказалось, что существует незначительная положительная связь между среднемесячной температурой воздуха и месячным доходом на данном маршруте трамвая.

По принципу своего построения коэффициент корреляции предназначен для измерения только *линейной* статистической связи между двумя переменными. При прямой линейной зависимости между x и y $r_{xy} \approx 1$, при прямой статистической связи $0 < r_{xy} < 1$. В случае обратной линейной зависимости между x и y $r_{xy} \approx -1$, а в остальных случаях обратной статистической связи $-1 < r_{xy} < 0$. При $r_{xy} = 0$ нужны дополнительные исследования.

Приблизительно определить значение коэффициента корреляции можно, анализируя диаграмму рассеяния. Чем теснее расположены точки относительно некоторой прямой, тем больше по абсолютной величине r_{xy} , и наоборот, чем более расплывчато облако точек на диаграмме рассеяния, тем ближе к нулю коэффициент корреляции (рис. 4.2).

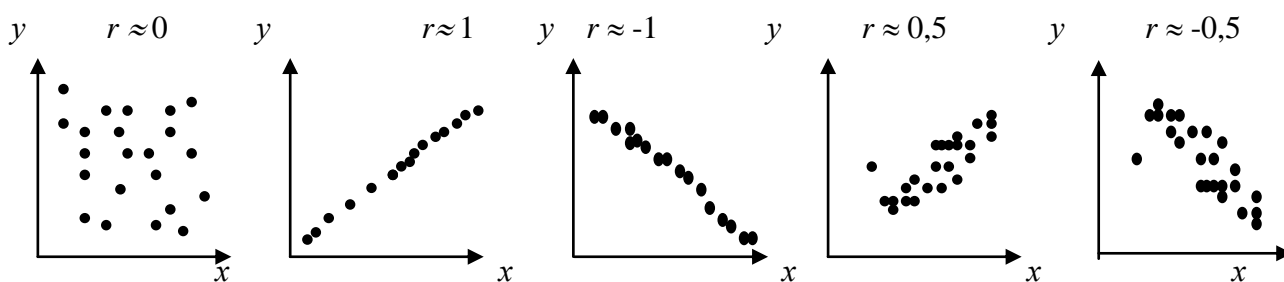


Рис. 4.2. Диаграмма рассеяния при различных значениях коэффициента корреляции

Величины x и y , для которых рассчитывается коэффициент корреляции, могут иметь произвольные средние и дисперсии. Исходные величины x и y можно умножать на произвольную константу одного знака и прибавлять к ним произвольные константы, не изменяя при этом величину коэффициента корреляции между ними.

Рассмотрим смежные вопросы, связанные с использованием коэффициента корреляции.

Довольно часто в исследовательской практике требуется определить дисперсию суммы или разности двух или более случайных величин. Выведем формулу для случая двух слагаемых. Имеем

$$\begin{aligned} s_{x+y}^2 &= \frac{\sum ((x+y) - \overline{(x+y)})^2}{n-1} = \frac{\sum ((x+y) - (\bar{x} + \bar{y}))^2}{n-1} = \frac{\sum ((x-\bar{x}) + (y-\bar{y}))^2}{n-1} = \\ &= \frac{\sum (x-\bar{x})^2}{n-1} + \frac{2\sum (x-\bar{x})(y-\bar{y})}{n-1} + \frac{\sum (y-\bar{y})^2}{n-1}. \end{aligned} \quad (4.4)$$

Из соотношения (4.4) видно, что первый и третий член правой части представляют собой дисперсии s_x^2 и s_y^2 , а второй член – удвоенную ковариацию s_{xy} . Таким образом,

$$s_{x+y}^2 = s_x^2 + s_y^2 + 2s_{xy}. \quad (4.5)$$

Учитывая, что $s_{xy} = r_{xy} s_x s_y$, получаем

$$s_{x+y}^2 = s_x^2 + s_y^2 + 2r_{xy} s_x s_y. \quad (4.6)$$

Если величины x и y некоррелированы, в частности если они независимы, то $r_{xy} = 0$ и соотношение (4.6) приобретает вид

$$s_{x+y}^2 = s_x^2 + s_y^2. \quad (4.7)$$

В том случае, когда $r_{xy} \neq 0$, анализ уравнения (4.6) показывает, что положительная корреляция между x и y приводит к увеличению дисперсии их суммы, а отрицательная – к уменьшению.

По аналогии с приведенными преобразованиями легко получить

$$s_{x-y}^2 = s_x^2 + s_y^2 - 2r_{xy} s_x s_y. \quad (4.8)$$

Полученные формулы (4.6) и (4.8) обобщаются на случай более чем двух переменных. Например, для случая трех переменных (4.6) обобщается следующим образом

$$s_{x+y+z}^2 = s_x^2 + s_y^2 + s_z^2 + 2r_{xy} s_x s_y + 2r_{xz} s_x s_z + 2r_{yz} s_y s_z. \quad (4.9)$$

Теперь рассмотрим вопрос, как влияют на величину коэффициента корреляции ошибки измерения тех признаков, между которыми оценивается взаимосвязь.

Пусть u и v – измеряемые величины, связанные с истинными значениями признаков x и y следующими соотношениями:

$$u = x + \delta, \quad (4.10)$$

$$v = y + \varepsilon. \quad (4.11)$$

Будем считать, что случайные ошибки δ и ε имеют средние, равные нулю, и некоррелированы с u и v . Тогда, $\bar{u} = \bar{x}$, $\bar{v} = \bar{y}$, а $s_u^2 = s_x^2 + s_\delta^2$, $s_v^2 = s_y^2 + s_\varepsilon^2$ в силу полученного соотношения (4.7).

С учетом вышесказанного легко показать, что $s_{uv} = s_{xy}$. И если теперь записать выражения для r_{xy} и r_{uv} и взять их отношение, то получим

$$r_{xy} = \frac{s_{xy}}{s_x s_y}; \quad r_{uv} = \frac{s_{uv}}{s_u s_v}; \quad \frac{r_{xy}}{r_{uv}} = \frac{s_{xy} s_u s_v}{s_x s_y s_{uv}} = \frac{s_u s_v}{s_x s_y}. \quad (4.12)$$

Ввиду того что $s_u > s_x > 0$ и $s_v > s_y > 0$, то и $s_u s_v > s_x s_y$, откуда следует что если при измерениях допускаются ошибки, то наблюдаемая корреляция меньше истинной, причем с увеличением числа наблюдений их расхождение не уменьшается.

Рассмотрим, как изменяется коэффициент корреляции в результате смешивания коррелированных пар с некоррелированными. Пусть у нас имеется n_1 пар наблюдений над x и y , которые дали коэффициент корреляции r_1 . И пусть к этому материалу добавлено n_2 пар, имеющих средние и квадратичные отклонения, равные тем, которые были вычислены для первых n_1 пар, но с

равным нулю коэффициентом корреляции. При таких условиях коэффициент корреляции объединенной совокупности будет равен

$$r_2 = \frac{\sum (x - \bar{x})(y - \bar{y})}{(n_1 + n_2 - 1)s_x s_y}.$$

Откуда

$$\frac{r_2}{r_1} = \frac{n_1 - 1}{n_1 + n_2 - 1}. \quad (4.13)$$

Таким образом видно, что в результате смешения получаемый коэффициент корреляции будет меньше.

4.3. ИНТЕРПРЕТАЦИЯ КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ

По поводу интерпретации r_{xy} следует сделать два важных замечания. При выводе формулы для коэффициента корреляции никаких специальных предположений о характере совместного двумерного распределения величин x и y не делалось. Однако выводы о пределах изменения от -1 до $+1$ для r_{xy} , а также выражения «сильная связь», «умеренная связь» и т. д. справедливы только в рамках определенной статистической модели. В частности, сделанные выводы справедливы в том случае, если совместное распределение x и y представляет собой нормальную двумерную поверхность – симметричную колоколообразную поверхность.

Нарушение этого предположения не препятствует вычислению r_{xy} , однако его интерпретация существенно меняется. Так, если частотные распределения x и y существенно скошены в разных направлениях, то даже при максимально возможной линейной связи между x и y r_{xy} не будет превышать по абсолютной величине $0,6-0,7$.

Такая зависимость максимальной величины r_{xy} от характера распределения x и y приводит к трудностям при интерпретации получаемых конкретных значений r_{xy} . В самом деле, что означает величина $r_{xy} = 0,6$? Максимально возможную линейную связь при положительной и отрицательной асимметрии распределений x и y или умеренную связь этих переменных при совместном распределении, подобном двумерному нормальному? Ответы на эти вопросы и, следовательно, дальнейшая интерпретация r_{xy} могут быть получены из качественного анализа диаграмм рассеяния и гистограмм распределения x и y . Второе замечание по поводу интерпретации коэффициентов корреляции связано со значениями этих коэффициентов, близкими к нулю.

Рассмотрим диаграммы рассеяния, приведенные на рис. 4.3.

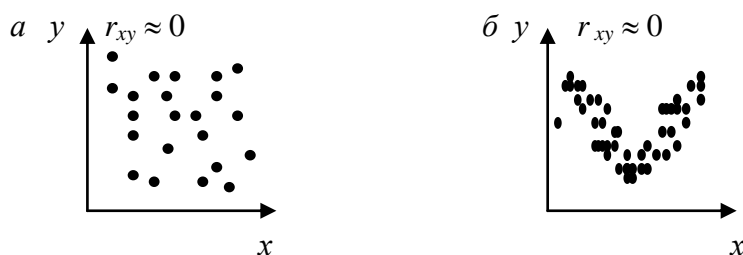


Рис. 4.3. Диаграммы рассеяния для $r_{xy} \approx 0$

Ранее уже подчеркивалось, что r_{xy} измеряет тесноту только линейной связи между переменными x и y . Поэтому для диаграммы рассеяния, приведенной на рис. 4.3, а, $r_{xy} \approx 0$. Однако $r_{xy} \approx 0$ и для диаграммы рассеяния 4.3, б, несмотря на то, что в этом случае имеет место почти функциональная зависимость. Происходит это из-за нелинейной связи между x и y .

Этот пример должен служить предостережением против опрометчивого вывода о том, что равенство нулю коэффициента корреляции между переменными x и y всегда свидетельствует об отсутствии статистической связи между ними.

Поэтому для исключения возможных ошибок содержательной интерпретации коэффициента корреляции должен сопутствовать визуальный анализ диаграммы рассеяния. Но даже в тех случаях, когда исходные предпосылки корреляционного анализа выполняются, исследователя при объяснении смысла полученных значений r_{xy} подстерегают опасности, связанные с так называемой ложной корреляцией. Проблемы, которые при этом возникают, проиллюстрируем на примере расчетов для двух искусственных рядов, взятых из таблицы случайных чисел:

x	0	4	2	6	6	2	9	8	5	3
y	8	9	1	0	5	2	4	6	7	1

Для этих двух рядов, как и следовало ожидать, коэффициент корреляции близок к нулю: $r_{xy} = 0,006$.

Пусть теперь имеется еще один ряд z : 7, 2,5, 7, 2, 3, 9, 2, 2, 1, по отношению, к которому определены два ряда $x' = x/z$ и $y' = y/z$. Таким образом, по отношению к x' и y' z выступает в виде общей причины. Под действием этой общей причины ряды x' и y' принимают следующий вид:

x'	0,00	2,00	0,40	0,86	3,00	0,67	1,00	4,00	2,50	3,00
y'	1,14	4,50	0,20	0,00	2,50	0,67	0,44	3,00	3,50	1,00

Коэффициент корреляции в этом случае $r_{xy} = 0,583$.

Таким образом, несмотря на отсутствие статистической связи между x и y , действие общей причины приводит к $r_{xy} \neq 0$, указывающему на умеренную корреляционную связь, что может послужить основой для неправильного истолкования взаимосвязи между x и y , когда механизм связи неизвестен.

Другой возможный вариант возникновения ложной корреляции связан с появлением в совокупности исходных данных аномальных значений либо за счет неоднородности анализируемого материала, либо за счет ошибок при регистрации данных. Предположим, что случайные ряды x и y , которые были использованы ранее, получены из показаний какого-то прибора путем деления этих показателей на 5, для того чтобы перейти, например, от условных единиц к истинным значениям. При этом для одной пары значений по ошибке такой переход не был сделан. В этом случае ряды x и y будут выглядеть следующим образом:

x	0	4	2	6	6	2	9	8	25	3
y	8	9	1	0	5	2	4	6	35	1.

Одной такой ошибки достаточно, для того чтобы коэффициент корреляции стал равным 0,874, т. е. значение r_{xy} в этом случае указывает на довольно тесную связь, что в свою очередь может привести к неправильным содержательным выводам.

Наконец, ложная корреляция может возникнуть за счет того, что при анализе взаимосвязи исследователь произвольно выбирает из всего множества экспериментальных данных некоторый их набор, в котором пары величин изменяются однонаправленно в силу совпадения случайных обстоятельств. Например, в приводимом ниже примере значения x и y взяты из таблицы случайных чисел, но не подряд, а по определенному правилу. В результате имеем два ряда:

x	3	6	7	1	2	4	0	2	9	8
y	7	2	1	9	7	5	6	3	0	4

для которых $r_{xy} = -0,780$.

Таким образом, приведенные примеры еще раз подтверждают соображения, высказанные при обсуждении общих проблем корреляции, о необходимости подробного логико-теоретического анализа явлений, одной из оценок которых является коэффициент корреляции. Только такой анализ может помочь исследователю избежать ошибок при интерпретации полученных коэффициентов корреляции.

5. РЕГРЕССИОННЫЙ АНАЛИЗ

Довольно часто в практике исследовательской работы имеет место ситуация, когда важнейшие переменные, описывающие некоторый процесс, известны заранее, но модель процесса еще неизвестна. В этом случае возможны разные подходы. Одним из них является построение эмпирических моделей.

Построение эмпирических моделей предполагает проведение экспериментов или наблюдений для сбора опытных данных, выбор одной определенной модели из некоторого множества возможных, вычисление коэффициентов модели («подгонку») и оценку полученных результатов.

Рассмотрим только этап, связанный с вычислением коэффициентов модели. Вообще говоря, эмпирические модели могут быть как линейными, так и нелинейными. Причем термин «линейность» может относиться как к независимым переменным модели, так и к ее коэффициентам (параметрам). Рассмотрим линейную модель с линейной комбинацией параметров [10].

5.1. МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

Метод наименьших квадратов разработан знаменитыми математиками К. Гауссом и А. Лежандром. Рассмотрим этот метод на примере линейной модели

$$y^* = a + bx,$$

где x – независимая переменная, т. е. переменная, которую экспериментатор может изменить по своему усмотрению; y^* – зависимая переменная или отклик; a и b – коэффициенты (параметры). Из данных, приведенных в табл. 5.1 и рис. 5.1, видно, что такой моделью (уравнением прямой линии) может быть

описана зависимость между суммарными простоями транспорта на маршруте трамвая за месяц (в часах) и месячным доходом (в тыс. руб.). С другой стороны, из рис. 5.1 видно также, что реально наблюдаемые значения отклика y_i несколько отличаются от откликов y_i^* , соответствующих уравнению модели. И такое положение будет всегда, даже в тех случаях, когда зависимая и независимая переменные будут связаны строгой функциональной зависимостью. В этом случае отклонения эмпирических значений от теоретических связаны с погрешностями измерений, которые всегда имеют место.

Таблица 5.1

Суммарные простои на маршруте за месяц p , ч	Доход на маршруте за месяц Σ , тыс. руб.	Суммарные простои на маршруте за месяц p , ч	Доход на маршруте за месяц Σ , тыс. руб.
278,88	1930,278	81,88	2443,99
194,23	2117,094	81,53	2453,544
191,27	2642,654	158,50	2745,722
154,42	2663,458	208,98	2932,050
81,25	2532,026	336,45	2868,228
149,83	2512,664	597,85	2857,032

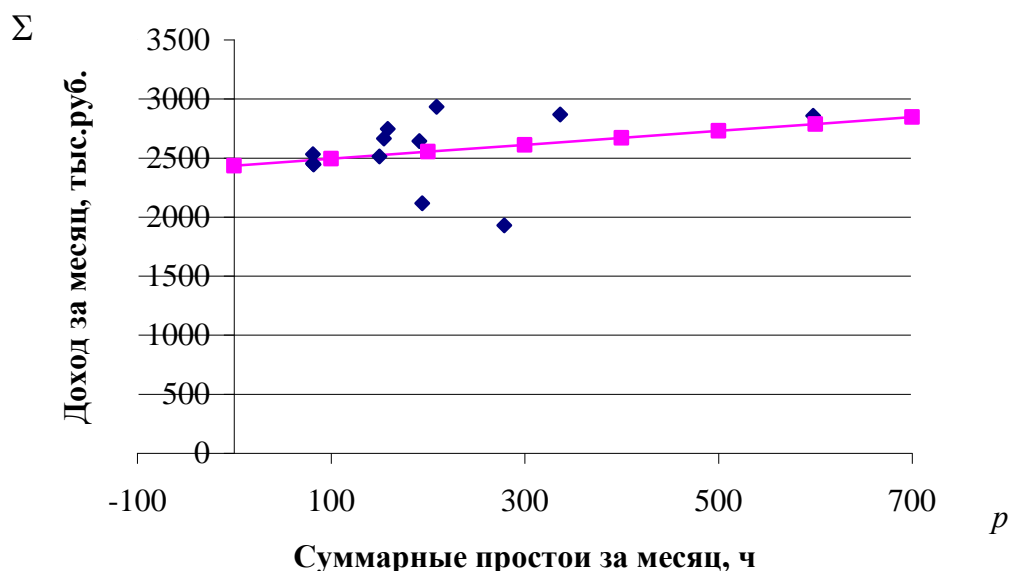


Рис. 5.1. Зависимость между суммарными простоями транспорта на маршруте за месяц (p , ч) и выручкой (Σ , тыс. руб.)

Итак, каждому значению независимой переменной в общем случае соответствует ошибка

$$\xi_i = y_i - y_i^*. \quad (5.1)$$

Естественно, что в зависимости от того, как будет проведена прямая, аппроксимирующая набор экспериментальных данных, величины ξ_i будут

различны. Именно для того, чтобы избежать субъективности при построении эмпирической модели, и был разработан метод наименьших квадратов, позволяющий однозначно определять параметры выбранной модели. В основе этого метода лежит критерий минимизации суммы квадратов ошибок, т. е. требование, чтобы

$$S = \sum_{i=1}^n \xi_i^2 = \sum_{i=1}^n (y_i - y_i^*)^2. \quad (5.2)$$

была минимальной.

Покажем, как используется метод наименьших квадратов на примере оценки параметров для уравнения $y^* = a + bx$.

Начнем с преобразования формулы (5.2):

$$S = \sum_{i=1}^n (y_i - (a + bx_i))^2 = \sum_{i=1}^n (y_i^2 - 2y_i(a + bx_i) + (a + bx_i)^2).$$

Опуская (для краткости) индекс i , получим:

$$S = \sum y^2 - 2a \sum y - 2b \sum xy + na^2 + 2ab \sum x + b^2 \sum x^2. \quad (5.3)$$

Нас интересует нахождение таких значений параметров a и b , которые обращают в минимум (5.3). Как известно, для этого необходимо найти частные производные $\partial S / \partial a$ и $\partial S / \partial b$ и приравнять их к нулю. В общем случае после такой процедуры необходимо убедиться, что в данной точке имеет место минимум (а не максимум или точка перегиба). Итак,

$$\frac{\partial S}{\partial a} = -2 \sum y + 2na + 2b \sum x = 2(na - \sum y + b \sum x) = 0, \quad (5.4)$$

$$\frac{\partial S}{\partial b} = -2 \sum xy + 2a \sum x + 2b \sum x^2 = 2(a \sum x - \sum xy + b \sum x^2) = 0. \quad (5.5)$$

Из (5.4) и (5.5) получаем систему двух уравнений, построенных по методу наименьших квадратов:

$$\begin{cases} \sum y = na + b \sum x; \\ \sum xy = a \sum x + b \sum x^2, \end{cases} \quad (5.6)$$

из которых определяются коэффициенты a и b .

Проиллюстрируем нахождение этих коэффициентов на данных, приведенных в табл. 5.1. Для этого вычислим сначала промежуточные величины: $\sum x = 2515,1$; $\sum x^2 = 757749,3$; $\sum xy = 6569650$. Имеем

$$\begin{cases} 30698,7 = 12a + 2515,1b; \\ 656965 = 2515,1a + 757749,3b. \end{cases}$$

Откуда $a = 2435,1$; $b = 0,6$ и модель, связывающая суммарные простои транспорта на маршруте трамвая за месяц (в часах) и выручку (в тыс. руб.), принимает вид $y^* = 2435,1 + 0,6x$.

В табл. 5.2 проведено сравнение между реальными данными и данными, полученными с использованием эмпирической модели, а также величины ошибок. Минимально возможная сумма квадратов ошибок $\sum_{i=1}^n \xi_i^2 = 914529$.

Таблица 5.2

y	y^*	ξ_i
1930,278	2598,948	668,670
2117,094	2549,202	432,108
2642,654	2547,463	-95,191
2663,458	2525,807	-137,651
2532,026	2482,807	-49,219
2512,664	2523,110	10,446
2443,990	2483,178	39,188
2453,544	2482,972	29,428
2745,722	2528,205	-217,517
2932,050	2557,870	-374,180
2868,228	2632,780	-235,448
2857,032	2786,397	-70,635

Рассмотрим уравнение (5.6) с точки зрения их связи с другими описательными статистиками. Разделив первое из уравнений на n , получим

$$\frac{\sum y}{n} = a + b \frac{\sum x}{n}; \quad \bar{y} = a + b\bar{x}. \quad (5.7)$$

Подставим значения для a из (5.7) во второе уравнение из (5.6):

$$\sum xy = (\bar{y} - b\bar{x})\sum x + b\sum x^2 = \bar{y}\sum x + b(\sum x^2 - \bar{x}\sum x).$$

Если первое слагаемое из правой части перенести в левую и разделить обе части уравнения на $n(n-1)$, то

$$\frac{\frac{\sum xy}{n} - \bar{y} \frac{\sum x}{n}}{n-1} = \frac{b \left(\frac{\sum x^2}{n} - \bar{x} \frac{\sum x}{n} \right)}{n-1},$$

откуда следует, что

$$s_{xy} = bs_x^2. \quad (5.8)$$

Из двух последних уравнений имеем

$$\begin{cases} a = \bar{y} - b\bar{x}, \\ b = \frac{s_{xy}}{s_x^2}. \end{cases} \quad (5.9)$$

Уравнение (5.8) определяет семейство параллельных прямых линий с угловым коэффициентом b , а уравнение (5.7) позволяет выбрать из этого семейства ту прямую, которая проходит через точку с координатами, равными средним значениям переменных.

Существует связь между коэффициентом b и коэффициентом корреляции r_{xy} . Умножим обе части второго уравнения в (5.9) на s_x/s_y :

$$b \frac{s_x}{s_y} = \frac{s_{xy}}{s_x^2} \cdot \frac{s_x}{s_y} = \frac{s_{xy}}{s_x s_y} = r_{xy}; \quad b = r_{xy} \frac{s_y}{s_x}. \quad (5.10)$$

В частности, если $s_y = s_x$ то $b = r_{xy}$. Учитывая, что r_{xy} изменяется в пределах +1 до -1 делаем вывод, что коэффициент b не превосходит по абсолютной величине отношение среднеквадратичных отклонений для y и x : s_y/s_x .

5.2. ВЫБОР ФОРМЫ ФУНКЦИОНАЛЬНОЙ ЗАВИСИМОСТИ

Пусть имеется ряд данных, представляющих одну зависимую и одну независимую переменную, и требуется определить функциональную связь между ними. Универсального способа решения этой задачи не существует. Иногда анализ графического изображения имеющихся данных, а также понимание механизма исследуемого процесса помогают выбрать вид аналитической зависимости. Особенно просто оценить вид функциональной зависимости, если экспериментальные данные укладываются или группируются относительно некоторой прямой. Поэтому, если удастся найти такое преобразование исходных величин, что вновь полученные значения укладываются на прямую, вид исходной функциональной зависимости определить относительно несложно.

Значительное число зависимостей, встречающихся в практике научных исследований в самых разных областях знания, может быть описано следующими уравнениями:

$$y = a + bx, \quad (5.11)$$

$$y = ab^x, \quad (5.12)$$

$$y = ax^b, \quad (5.13)$$

$$y = a + bx + cx^2, \quad (5.14)$$

$$y = a + \frac{b}{1 + cx}. \quad (5.15)$$

Уравнения (5.11)–(5.15) аппроксимируют выборочную совокупность. Параметры a , b , c , которые осуществляют наилучшую аппроксимацию, можно найти методом наименьших квадратов (см. раздел 5.1).

Уравнение (5.11) имеет линейный вид, а уравнения (5.12) и (5.13) легко привести к линейному виду, прологарифмировав их левые и правые части:

$$\lg y = \lg a + x \lg b \text{ или } Y = A + Bx; \quad (5.12a)$$

$$\lg y = \lg a + b \lg x \text{ или } Y = A + bX, \quad (5.13a)$$

где $Y = \lg y$, $A = \lg a$, $B = \lg b$, $X = \lg x$. Параметры уравнений (5.12a), (5.13a) можно определить, используя систему уравнений (5.6). При этом для (5.11) надо использовать массивы x , y , для (5.12a) – массивы x , Y , для (5.13a) – массивы X , Y .

Параметры a , b , c уравнений (5.14) и (5.15) можно найти с помощью сервисной функции *MS Excel* «Сервис – Поиск решения». Параметры a , b , c , которые осуществляют наилучшую аппроксимацию, можно найти методом наименьших квадратов (см. раздел 5.1).

Таким образом, если известен вид функциональной зависимости, то найти коэффициенты не трудно. Сложно определить какое из уравнений является

наиболее подходящим. Технику подбора подходящей аналитической зависимости продемонстрируем на данных, приведенных в табл. 5.3 (рис. 5.2).

Таблица 5.3

День недели (1 – понедельник и т.д.)	Дневной доход на маршруте трамвая, тыс. руб.
1	91,742
2	88,018
3	87,542
4	76,510
5	81,914
6	50,666
7	41,216

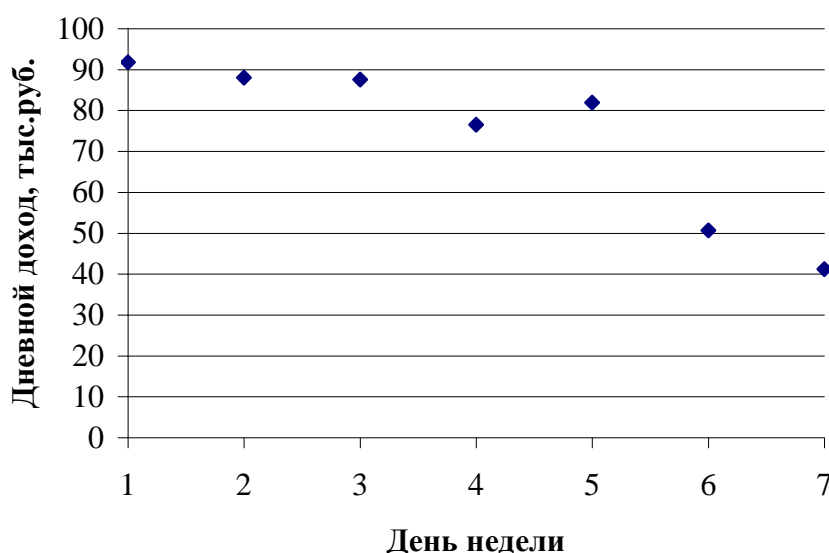


Рис. 5.2. Зависимость дневного дохода от дня недели

Попробуем аппроксимировать эти данные, используя уравнения (5.11)–5.13). Для уравнения (5.11) определили $a = 107,07$, $b = -8,28$. Получаем уравнение линейной аппроксимации $y = 107,07 - 8,28x$.

Для уравнения (5.12a) определили $A = 2,074$, $B = -0,055$. Отсюда $a = 10^A = 118,549$, $b = 10^B = 0,88$. После потенцирования (5.12a) получаем уравнение показательной аппроксимации $y = 118,55 \cdot (0,88)^x$.

Для уравнения (5.13a) определили $A = 2,035$, $b = -0,35$. Отсюда $a = 10^A = 108,35$. После потенцирования (5.13a) получаем уравнение степенной аппроксимации $y = 108,35,6x^{-0,35}$.

Графики трех аппроксимаций представлены на рис. 5.3.

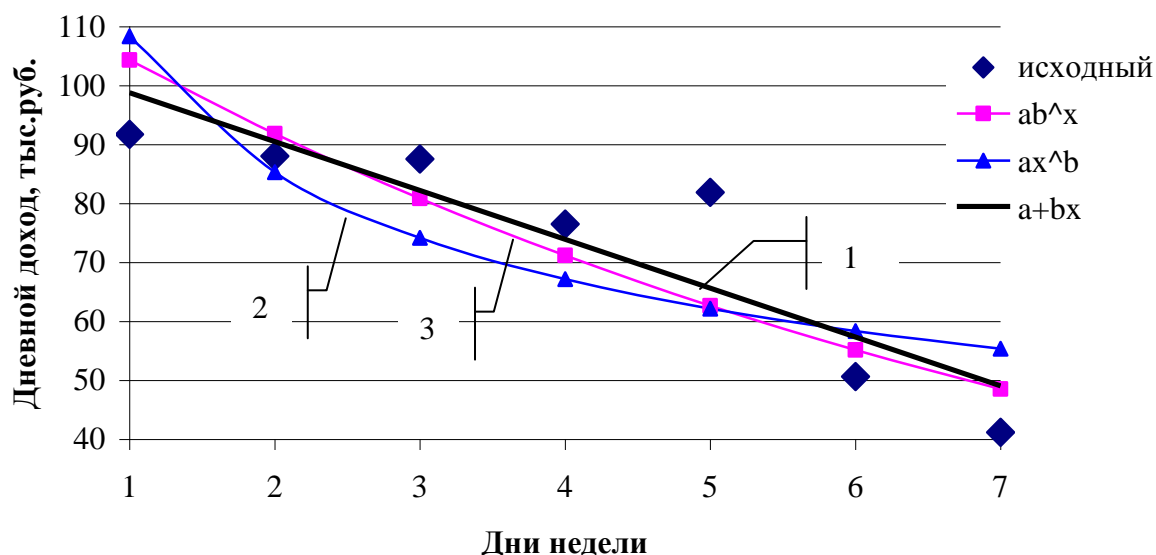


Рис. 5.3. Аппроксимации экспериментальных данных:
1 – линейная, 2 – показательная, 3 – степенная

Минимальная квадратичная невязка получилась для показательной аппроксимации $\sum_{i=1}^n \xi_i^2 = 691,82$. Однако эту невязку можно уменьшить, если

перейти от уравнения (5.12) к семейству

$$y = ab^{x+c} + d. \quad (5.126)$$

С помощью сервисной функции *MS Excel* «Сервис – Поиск решения» найдем решение этого уравнения. Функция квадратичной невязки для уравнения (5.126) имеет множество локальных минимумов. Наилучший результат при поиске глобальных минимумов зависит от выбора начальных значений параметров. При найденных $a = -11,48$, $b = 1,54$, $c = -3,30$, $d = 95,97$ (рис. 5.4) получили минимальную квадратичную невязку $\sum_{i=1}^n \xi_i^2 = 192,96$.

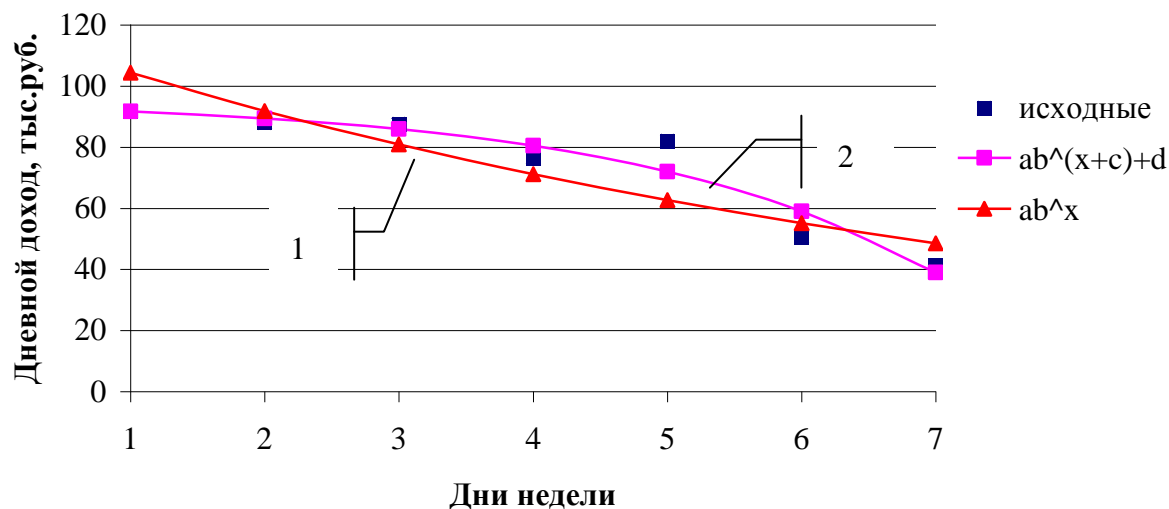


Рис. 5.4. Аппроксимации экспериментальных данных:
 $1 - y = ab^x$; $2 - y = ab^{x+c} + d$

5.3. ЛИНЕЙНЫЕ РЕГРЕССИОННЫЕ УРАВНЕНИЯ

Рассмотренный в разделе 5.1 метод наименьших квадратов представляет основной способ оценивания в регрессионном анализе, т. е. анализе, связанном с изучением зависимости среднего одной из случайных величин от другой или других случайных величин, связанных с ней.

Основы регрессионного анализа, так же как и корреляционного, были разработаны известным английским антропологом и метеорологом Ф. Гальтоном. Он же ввел в научный оборот и само это слово в применении к широкому классу задач. Будучи близким родственником Чарльза Дарвина, он решил оказать ему помощь в количественной оценке некоторых фактов, имеющих отношение к теории естественного отбора. Он провел обширное экспериментальное исследование, измеряя рост сыновей и отцов. Гальтон пришел к выводу, что у сыновей имеет место регрессия к среднему: не проявляется тенденция к наследованию сыновьями роста отцов, и у высоких отцов сыновья, как правило, ниже, а у низкорослых наоборот.

В разделе 5.2 было показано, что зависимость одной случайной величины от другой может описываться различными аналитическими выражениями. Рассмотрим простейший случай линейной зависимости и соответствующие ему линейные регрессионные уравнения. Будем считать, что каждому значению независимой переменной x_i соответствуют несколько значений зависимой переменной y . Таким образом, имеем двумерную выборку $\{(x_i, y_i)\}$, $i = \overline{1, n}$, для которой x_i могут повторяться.

Естественно, что уравнения, составленные по методу наименьших квадратов, всегда позволяют вычислить значения коэффициентов a и b даже в тех случаях, когда выбранное уравнение регрессии $y = a + bx + e$ не слишком удачно приближает данные. В значительной мере достоверность полученных оценок a и b зависит от некоторых предположений относительно поведения случайной ошибки e .

Обозначим через $\{x^j\}$, $j = \overline{1, k}$, множество различных значений x_i . Через y^j обозначим пересечение всех y_i с вертикальной прямой $x = x^j$ $y^j_{\text{выб}} = \{y_i: x_i = x^j\}$, $j = \overline{1, k}$. Множество $y^j_{\text{выб}}$ представляют собой выборки некоторых случайных величин y^j . Аналогично множества $e^j_{\text{выб}} = y^j_{\text{выб}} - a - bx^j$ являются выборками случайных величин e^j . При этом объединение всех $e^j_{\text{выб}}$ образуют выборку случайной ошибки e . Перечислим требования к случайной величине e .

1. Случайный характер: отдельные ошибки e^j представляют собой случайные величины.

2. Нулевое среднее: каждое отклонение e^j характеризуется нулевым математическим ожиданием и не зависит от значений x^j .

3. Гомоскедастичность: дисперсия каждого e^j , одинакова для всех j независимо от x^j .

4. Отсутствие взаимозависимости (автокорреляции) ошибок: значения ошибок отдельных наблюдений не коррелированы между собой.

5. Нормальность: требование, чтобы ошибка e имела нормальное распределение.

В следующих главах будут описаны приемы, с помощью которых можно проверить выполнение этих требований. Здесь же будем предполагать, что они выполняются, и попытаемся приближенно оценить точность предсказания y по x , так как именно предсказание является одной из главных целей регрессионного анализа.

Такая оценка будет осуществляться с использованием несмещенной дисперсии ошибки оценки s_e^2 . Имеем

$$s_e^2 = \frac{\sum_{i=1}^n (e_i - \bar{e})^2}{n-1} = \frac{\sum_{i=1}^n e_i^2}{n-1}, \quad (5.16)$$

так как согласно сделанному предположению $\bar{e} = 0$ и все $\bar{e}^j = 0$.

Выразим s_e^2 через s_y^2 и r_{xy} , что позволит выявить некоторые связи между корреляцией и предсказанием. Итак,

$$s_e^2 = \frac{\sum_{i=1}^n e_i^2}{n-1} = \frac{\sum_{i=1}^n (y_i - a - bx_i)^2}{n-1}. \quad (5.17)$$

С учетом (5.17)

$$\begin{aligned} s_e^2 &= \frac{\sum (y_i - bx_i + b\bar{x} - \bar{y})^2}{n-1} = \frac{\sum ((y_i - \bar{y}) - b(x_i - \bar{x}))^2}{n-1} = \\ &= \frac{\sum (y_i - \bar{y})^2}{n-1} - \frac{2b \sum (x_i - \bar{x})(y_i - \bar{y})}{n-1} + \frac{b^2 \sum (x_i - \bar{x})^2}{n-1} = s_y^2 - 2bs_{xy} + b^2 s_x^2. \end{aligned}$$

Воспользовавшись соотношением (5.10), получим

$$s_e^2 = s_y^2 - 2r_{xy} \frac{s_y}{s_x} s_{xy} + r_{xy}^2 \frac{s_y^2}{s_x^2} s_x^2 = s_y^2 - 2r_{xy}^2 s_y^2 = s_y^2 (1 - r_{xy}^2). \quad (5.18)$$

Положительное значение корня квадратного из дисперсии ошибки оценки носит название стандартной ошибки:

$$s_e = s_y \sqrt{1 - r_{xy}^2}. \quad (5.19)$$

Эта ошибка оценки может быть использована для приближенного определения тех пределов в окрестностях предсказанного значения \hat{y} , внутрь которых попадает фактическое значение y . Если считать, что выполняется требование нормальности распределения ошибки, то можно использовать стандартную ошибку оценки для построения полосы относительно эмпирической линии регрессии, в которую попадают с определенной степенью вероятности истинные значения y при заданных x (рис. 5.5).

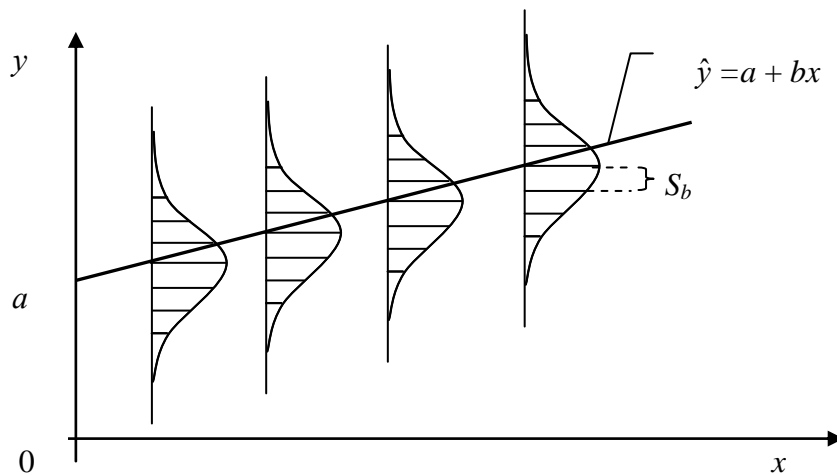


Рис. 5.5. Плотности распределений ошибки оценки e^j для приближенного значения \hat{y}

Линейный регрессионный анализ естественным образом обобщается на случай, когда зависимая переменная зависит не от одной, а от нескольких независимых переменных [12].

В табл. 5.4 приведены значения дневного дохода (y , тыс. руб.), среднедневной температуры воздуха (x_1 , °C) и простои транспорта (x_2 , ч) в течение зимнего месяца для одного маршрута троллейбуса.

Таблица 5.4

y	x_1	x_2	y	x_1	x_2	y	x_1	x_2	y	x_1	x_2
76,09	-22	43,07	69,87	-23	31,68	68,68	-7	43,12	43,51	-27	0,33
74,07	-23	35,48	70,07	-23	35,27	72,18	-6	31,92	34,55	-20	0,40
67,42	-24	31,7	76,23	-18	35,52	78,09	-13	40,48	76,20	-21	30,00
68,71	-14	37,52	72,77	-14	34,73	77,81	-23	34,70	71,62	-24	44,68
66,91	-19	35,72	67,34	-12	30,07	82,70	-14	36,15	80,47	-17	36,22
46,42	-17	5,40	45,98	-13	6,28	46,48	-29	13,97	68,38	-5	37,27
35,20	-21	1,05	42,43	-10	1,18	39,41	-31	1,47	41,66	-11	0,88

Необходимо построить уравнение регрессии, связывающее y с x_1 и x_2 : $y = a + bx_1 + cx_2$.

Неизвестные коэффициенты найдем из системы уравнений, составленной по методу наименьших квадратов. Имеем

$$\sum y = an + b\sum x_1 + c\sum x_2,$$

$$\sum x_1 y = a\sum x_1 + b\sum x_1^2 + c\sum x_1 x_2,$$

$$\sum x_2 y = a\sum x_2 + b\sum x_1 x_2 + c\sum x_2^2.$$

Используя статистическую функцию *MS Excel* ЛИНЕЙН(y , x , конст, статистика), получаем $a = 39,92$; $b = 0,02$; $c = 0,91$, и уравнение регрессии имеет вид

$$y = 39,92 + 0,02x_1 + 0,91x_2. \quad (5.20)$$

Если бы имелись только данные, связывающие дневной доход либо со среднедневной температурой воздуха, либо с простоями транспорта, то в этом случае регрессионные уравнения имели бы вид

$$y = 70,23 + 0,41x_1, \quad (5.21)$$

$$y = 39,62 + ,91x_2. \quad (5.22)$$

Стандартные (среднеквадратические) ошибки оценки, соответствующие уравнениям (5.20) – (5.22), равны $s_e = 5,17$, $s_{e'} = 15,28$, $s_{e''} = 5,18$.

Видно, что стандартные ошибки оценки при использовании двух независимых переменных меньше, чем при использовании только одной независимой переменной. Это следствие более общего результата, состоящего в том, что оценки, полученные по уравнению множественной регрессии, являются не менее надежными, чем оценки, полученные по уравнениям регрессии, включающим только одну независимую переменную. Как правило, с помощью уравнений множественной регрессии получается существенно лучшее приближение к экспериментальным данным, чем при учете только отдельных независимых переменных. Кроме того, данный пример показал, что наблюдается слабая зависимость дневного дохода на маршруте троллейбуса от среднедневной температуры воздуха.

На этом примере видно, что одновременный учет нескольких факторов, связанных с интересующей нас величиной, позволяет построить модель, которая более точно описывает имеющиеся данные и лучше прогнозирует зависимую переменную. Поэтому множественный регрессионный анализ довольно широко используется для построения уравнений, описывающих изучаемый процесс или явление.

Естественно, что интересующая нас величина может зависеть от большого числа факторов. Даже в тех случаях, когда любой из этих факторов поддается измерению, возникает вопрос, какие факторы имеет смысл учитывать при построении эмпирической модели в виде линейного уравнения регрессии, т. е. как из нескольких возможных моделей выбрать наиболее подходящую. Неразумно, даже имея в распоряжении вычислительную машину, рассчитывать все возможные уравнения регрессии, получаемые путем добавления или исключения из модели отдельных переменных.

Вообще говоря, процедура оценивания наиболее эффективна при правильно спланированном эксперименте, но рассмотрения требуют и те случаи, когда мы располагаем опытными данными, полученными в неспланированном заранее эксперименте. В этом случае нужно пользоваться правилами, которые позволяют выбрать наиболее подходящую модель с возможными наименьшими затратами сил и средств. Основные из них:

1. Модель – регрессионное уравнение – должна содержать минимальное число коэффициентов, а следовательно, и переменных, обеспечивающих приемлемую ошибку.

2. Модель должна иметь простейшую форму, совместимую с этой ошибкой. Например, при возможности нужно ограничиваться только первыми степенями переменных, не вводя квадратов и более высоких степеней.

3. Желательно, чтобы форма модели имела под собой содержательное

обоснование. Например, для системы автоматического регулирования при апериодическом переходном процессе в начальный период зависимость выходного сигнала от входного осуществляется по экспоненте, поэтому при изучении этого процесса должна быть выбрана соответствующая функциональная зависимость и т. д.

4. Сумма квадратов отклонений между эмпирическими и предсказанными по данному уравнению значениями должна быть минимальна.

Одной из лучших процедур, позволяющих осуществлять отбор наиболее подходящих моделей и учитывающих некоторые из приведенных выше правил, является шаговый регрессионный анализ. Сущность этого метода состоит в том, что отдельные переменные последовательно включаются в первоначальную модель, и на каждом этапе проводится проверка того, приводит ли добавление переменной к существенному или статистически значимому приближению предсказанных значений к эмпирическим данным. Такая проверка осуществляется с помощью дисперсионного анализа.

Необходимо также отметить, что уже для трех и более независимых переменных шаговый регрессионный анализ должен проводиться с использованием стандартных пакетов прикладных программ.

6. СТАТИСТИЧЕСКОЕ ОЦЕНИВАНИЕ И ПРОВЕРКА ГИПОТЕЗ

Различные статистические показатели вычисляются на основе некоторых ограниченных по объему совокупностей экспериментальных данных. Сбор и последующий анализ этих данных направлен на проверку определенных предположений относительно возможных механизмов изучаемых процессов или явлений, выдвигаемых исследователем при постановке задачи. Несмотря на то, что число объектов или процессов в каждом конкретном исследовании, как правило, сравнительно невелико, получаемые в результате анализа статистики должны достаточно хорошо описывать все реально существующее или мыслимое множество объектов или процессов, представляющее интерес в данном исследовании. Для этого делаются некоторые предположения о том, как вычисленные на основе экспериментальных данных показатели соотносятся с параметрами тех распределений, которые могли бы быть получены, если бы в нашем распоряжении оказалось все множество интересующих нас объектов. Это и составляет основную задачу статистического оценивания [10]. Решение этой задачи представляет собой главную часть любого анализа экспериментальных данных и тесно связано с использованием ряда теоретических распределений некоторых специальных переменных, которые и рассмотрим.

6.1. ТЕОРЕТИЧЕСКИЕ РАСПРЕДЕЛЕНИЯ, ИСПОЛЬЗУЕМЫЕ В СТАТИСТИЧЕСКИХ ВЫВОДАХ

Основное распределение в теории статистического оценивания – это нормальное распределение. Его свойства были подробно рассмотрены в предыдущих разделах, поэтому приведем только формулу для плотности этого распределения:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (6.1)$$

Широкое использование в статистических выводах нормального распределения имеет под собой как определенное эмпирическое, так и теоретическое обоснование. Во-первых, многочисленные примеры построения гистограмм и аппроксимация их непрерывными кривыми для экспериментальных данных самой различной природы показывают, что во многих случаях нормальное распределение действительно является довольно точным представлением таких данных. Во-вторых, доказано, что благодаря усреднению некоторого числа отдельных значений, которые сами по себе имеют произвольное распределение, получаемое распределение в среднем оказывается близким к нормальному. И наконец, нормальное распределение имеет ряд определенных математических достоинств, позволяющих существенно облегчить доказательства некоторых основополагающих моментов теории статистического вывода.

Следует четко представлять, что нормальное распределение – это чисто математический инструмент. И совсем необязательно, чтобы реальные экспериментальные данные точно описывались этим распределением, хотя во многих случаях, допуская небольшую ошибку, это распределение может быть принято в качестве модели для реальных данных.

Наряду с одномерным важную роль в статистике, особенно в теории корреляции, играет двумерное нормальное распределение:

$$f(x, y) = \frac{1}{\sqrt{2\pi}\sigma_x\sigma_y} e^{-\frac{1}{2}\left(\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right)}, \quad (6.2)$$

которое представляет собой колоколообразную поверхность в трехмерном пространстве (рис. 6.1) и обладает рядом характерных особенностей.

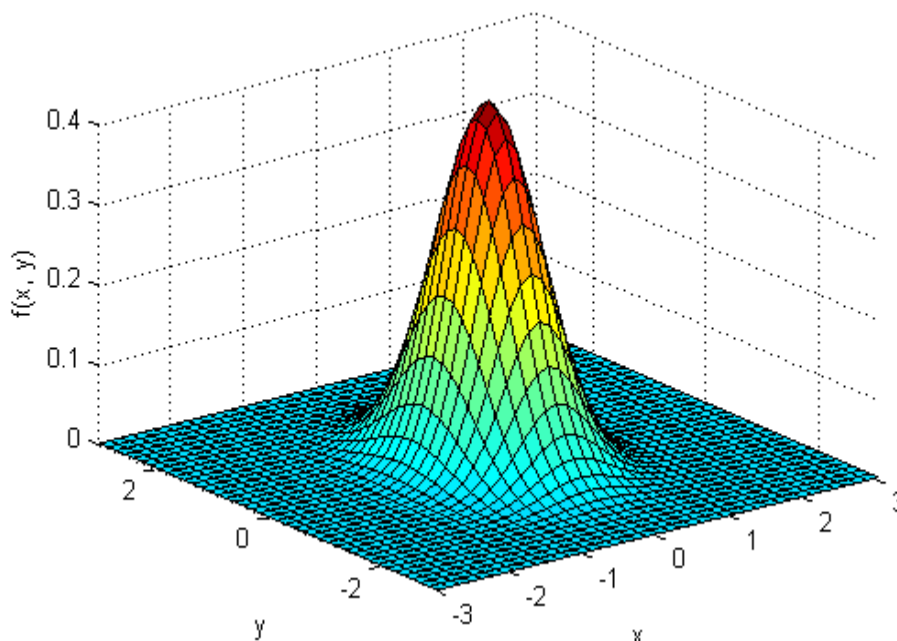


Рис. 6.1. Двумерное нормальное распределение

Во-первых, если взять только значения переменной x без учета значений переменной y , которым они соответствуют (предельное или частное), то распределение этих x является нормальным. То же справедливо и для значений y . Во-вторых, если зафиксировать какое-то значение переменной x , то все соответствующие ему значения переменной y (условное распределение) имеют также нормальное распределение (аналогично для y). И наконец, средние значения y для каждого отдельного значения x (условные средние) образуют прямую линию.

Дальнейшим обобщением нормального распределения является n -мерное нормальное распределение, широко используемое в многомерном статистическом анализе.

6.1.1. Распределение хи-квадрат

Другим играющим важную роль в статистических выводах распределением является распределение χ^2 (хи-квадрат), введенное в статистическую практику К. Пирсоном [13].

Возьмем выборочную совокупность нормально распределенной величины z с нулевым средним и единичным стандартным отклонением. Возведем элементы этого множества в квадрат. Получим выборку новой случайной величины $z^2 = \chi_1^2$. По выборке можно построить гистограмму и аппроксимировать (сгладить) ее непрерывной кривой. Путем умножения на константу выберем масштаб по оси ординат так, чтобы площадь под графиком равнялась единице. Получим кривую выборочного распределения хи-квадрат с одной степенью свободы. Одна степень свободы указывает, что для получения χ_1^2 использовались квадраты одной случайной величины.

Выберем случайно из исходной совокупности по два значения, возведем их в квадрат и сложим. Получим величину $\chi_2^2 = z_1^2 + z_2^2$ (хи-квадрат с двумя степенями свободы). Аналогично можно получить кривые выборочного распределения хи-квадрат для сумм произвольного числа квадратов стандартизованных величин, взятых из совокупности с нормальным распределением. Аналитическое выражение для кривой плотности распределения χ_n^2 :

$$f_{\chi_n^2}(z) = \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} z^{\frac{n}{2}-1} e^{-\frac{z}{2}},$$

где z – значение случайной величины χ_n^2 ; n – число степеней свободы; $\Gamma(n/2)$ – гамма-функция. Функции плотности распределения $f_{\chi_n^2}(z)$ табулированы.

На рис. 6.2 приведены плотности распределения случайных величин χ_n^2 для различных степеней свободы n .

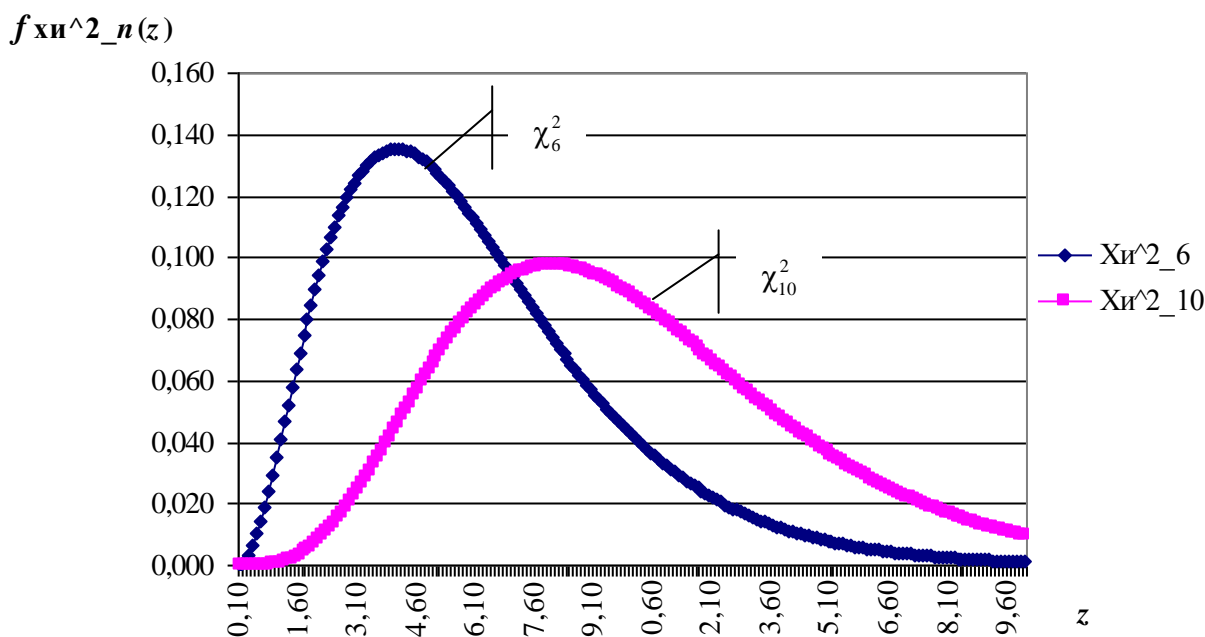


Рис. 6.2. Плотности распределения семейства χ^2_n

Как видно из рис. 6.2, свойства кривой χ^2_n существенно зависят от n . Семейство распределений хи-квадрат характеризуется следующими свойствами:

1. Среднее значение распределения хи-квадрат с n степенями свободы равно n .
2. Мода χ^2_n находится в точке $n - 2$ для $n \geq 2$.
3. Дисперсия χ^2_n равна $2n$.
4. Асимметрия χ^2_n равна $\sqrt{\frac{8}{n}}$.
5. Для больших n величин χ^2_n стремится к нормальному распределению со средним n и стандартным отклонением $\sqrt{2n}$.
6. Сумма двух независимых χ^2 -распределений $\chi^2_{n_1} + \chi^2_{n_2}$ является величиной того же семейства с $n_1 + n_2$ степенями свободы.

В [10] приведена таблица со значениями χ^2 для различных степеней свободы n , соответствующими определенным долям площади под кривой распределения. Например, что означает число 11,07 для $\alpha = 0,05$ и $n = 5$? Это число показывает, что сумма квадратов пяти значений z , случайно выбранных из стандартизованного нормального распределения, в среднем в пяти случаях из ста будет превышать 11,07. В остальных 95 случаях она будет меньше. Другими словами, вероятность получить значение χ^2_5 между 0 и 11,07 равна 0,95. Аналогичные табличные значения можно получить с помощью функции *Microsoft Excel* ХИ2РАСП($x; n$), которая показывает вероятность $P(\chi^2_n > x)$. В частности, ХИ2РАСП(11,07; 5) = 0,05.

6.1.2. F-распределение

Распределение Фишера (F -распределение) [13] получается следующим образом:

$$F_{n_1 n_2} = \frac{\chi_{n_1}^2 / n_1}{\chi_{n_2}^2 / n_2}, \quad (6.3)$$

где $\chi_{n_1}^2$ и $\chi_{n_2}^2$ – случайные независимые величины, имеющие распределение хи-квадрат; n_1, n_2 – степени свободы.

Можно построить множество F -отношений для разных пар (n_1, n_2) . Графики кривых F -распределения для разных чисел степеней свободы приведены на рис. 6.3.

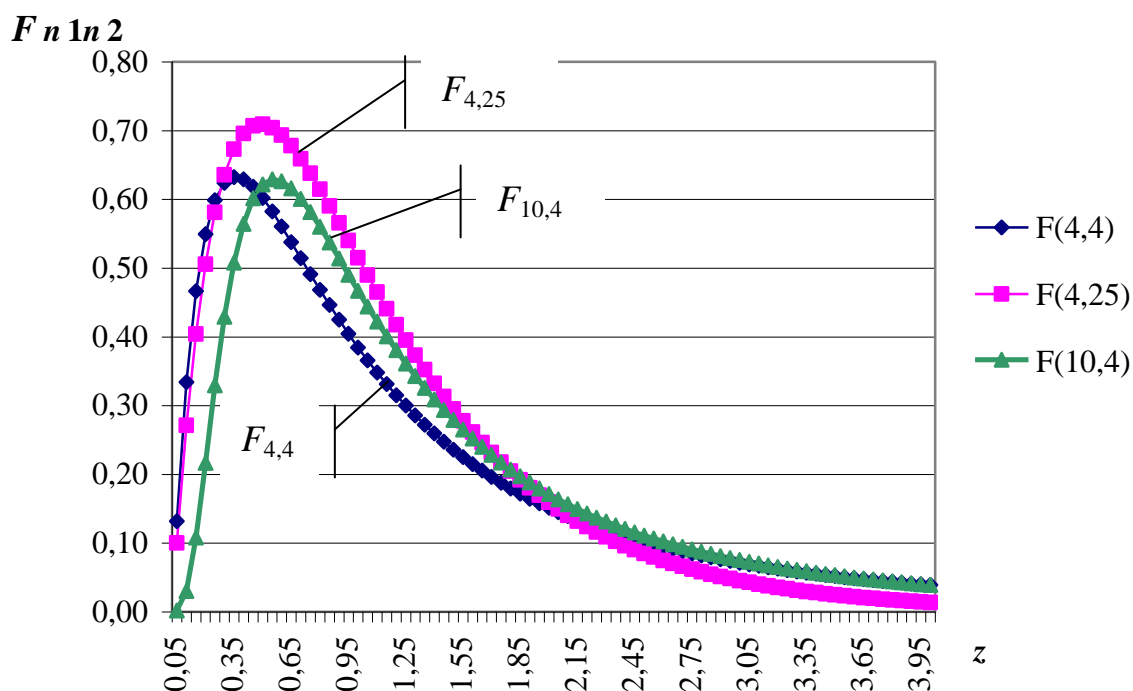


Рис. 6.3. F -распределение с (n_1, n_2) степенями свободы

F -распределение характеризуется унимодальностью, положительной асимметрией, имеет медиану меньше 1 и среднее, равное $n_1/(n_2 - 2)$ для $n_2 > 2$. В [10] приведена таблица F -распределения, в которой каждой паре степеней свободы соответствуют два числа для уровня значимости $\alpha=0,05$ и $\alpha=0,01$. Эти числа представляют собой граничные значения, слева от которых лежит соответственно 95 и 99 % от общей площади под кривой F -распределения. Аналогичные табличные значения можно получить с помощью функций *Microsoft Excel* $\text{FRасп}(x; n_1; n_2)$, которая вычисляет $P(x < F_{n_1 n_2})$. Например, $\text{FRасп}(3,327; 5; 10) = 0,05$, $\text{FRасп}(5,576; 5; 10) = 0,01$.

6.1.3. Распределение Стьюдента

Если в нашем распоряжении есть стандартизованное распределение z и распределение χ_n^2 то, как показал Стьюдент (псевдоним крупного английского статистика В. Госсета), величина

$$t = \frac{z}{\sqrt{\chi_n^2/n}} \quad (6.4)$$

имеет t -распределение [13]. Они образуют однопараметрическое семейство t -распределений (рис. 6.4).

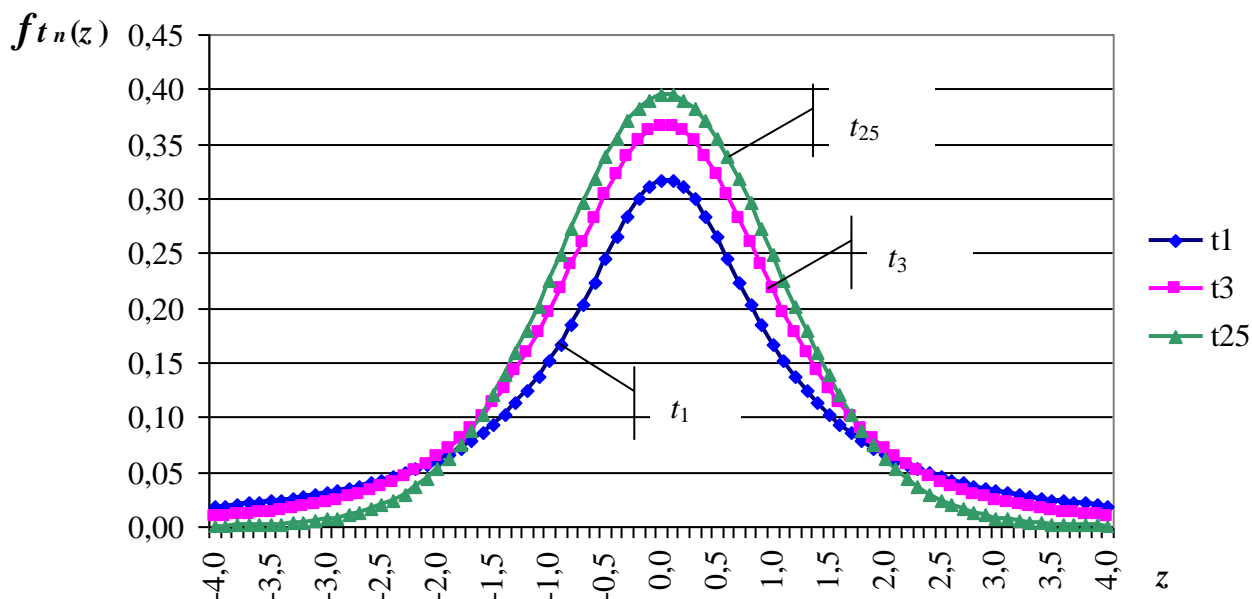


Рис. 6.4. Семейство t -распределений

Распределения симметричны, унимодальны, имеют нулевое среднее и дисперсию, равную $n/(n-2)$. По мере роста n распределение начинает быстро приближаться к нормальному. В [10] приведена таблица t -распределения для различных значений n .

Например, числа 2,365 и 3,499 для числа степеней свободы $n = 7$ означают, что только в пяти и соответственно одном случае из ста (для уровня значимости $\alpha=0,05$ и $\alpha=0,01$ соответственно) реализация t_n -распределения по абсолютной величине превысит эти числа. Или, другими словами, что в диапазонах $[-2,365; +2,365]$ и $[-3,499; +3,499]$ заключено соответственно 95 и 99 % площади под кривой t -распределения.

Аналогичные табличные значения можно получить с помощью функций *Microsoft Excel* СтюдРасп($x; n; \text{хвосты}$). Если «хвосты» = 2, то функция возвращает $P(X < -x \text{ или } X > x) = P(|X| > x)$, т.е. вычисляются два «хвоста» распределения. Если «хвосты» = 1, то функция возвращает $P(X > x)$, т.е. вычисляется только один «хвост» распределения. Например, СтюдРасп(2,365; 7; 2) = 0,05, СтюдРасп(3,499; 7; 2) = 0,01.

Распределение Стьюдента играет в теоретической и прикладной статистике почти такую же важную роль, как и нормальное распределение.

6.1.4. Биномиальное распределение

Одним из известных распределений дискретной случайной величины является биномиальное распределение. Оно является двухпараметрическим и

задается двумя параметрами: n – число испытаний, p – вероятность успеха в одном испытании. Обозначим биномиально распределенную СВ $X_{n,p}$ – количество успешных испытаний из общего количества испытаний. Тогда

$$P(X_{n,p}) = C_n^m p^m q^{n-m},$$

где $0 < p < 1$, $q = 1 - p$, $m = 0, 1, \dots, n$.

Плотность распределения величины является симметричной столбчатой диаграммой.

$$M(X_{n,p}) = np.$$

$$D(X_{n,p}) = npq.$$

6.1.5. Распределение Пуассона

Распределение Пуассона [13] так же, как и биномиальное описывает дискретные случайные величины.

Рассмотрим дискретную случайную величину X , которая может принимать только целые, неотрицательные значения: $0, 1, 2, \dots, m, \dots$, причем значения этой величины могут быть сколь угодно большими.

Случайная величина X распределена по закону Пуассона, если вероятность того, что она примет определенное значение m , выражается формулой

$$P_m = \frac{a^m}{m!} e^{-a}, \quad (6.5)$$

где a – некоторая положительная величина, называемая параметром закона Пуассона.

Закон распределения величины X имеет вид, приведенный ниже.

x	0	1	2	...	m	...
$P(X = x)$	e^{-a}	$\frac{a}{1!} e^{-a}$	$\frac{a^2}{2!} e^{-a}$...	$\frac{a^m}{m!} e^{-a}$...

Зададим параметру a некоторые численные значения и определим вероятности $P_m = P(X = m)$ для различных значений m по формуле (6.5). В результате этих действий получим данные рядов распределения. На их основе построим многоугольники распределения случайной величины X , распределенной по закону Пуассона (рис. 6.5). Из рис. 6.5 видно, что в зависимости от параметра a многоугольники распределения имеют существенные различия и по форме похожи на другие известные законы распределения случайной величины.

Математическое ожидание распределения Пуассона равно:

$$\begin{aligned} m_x = M[X] &= \sum_{m=0}^{\infty} m P_m = \sum_{m=0}^{\infty} m \frac{a^m}{m!} e^{-a} = \sum_{m=1}^{\infty} \frac{a^m}{(m-1)!} e^{-a} = \\ &= a e^{-a} \sum_{m=1}^{\infty} \frac{a^{m-1}}{(m-1)!} = a e^{-a} \sum_{n=0}^{\infty} \frac{a^n}{n!} = a e^{-a} e^a = a. \end{aligned} \quad (6.6)$$

Последняя сумма является разложением величины e^a в ряд Тейлора. Таким образом, параметр a представляет собой не что иное, как математическое ожидание случайной величины X .

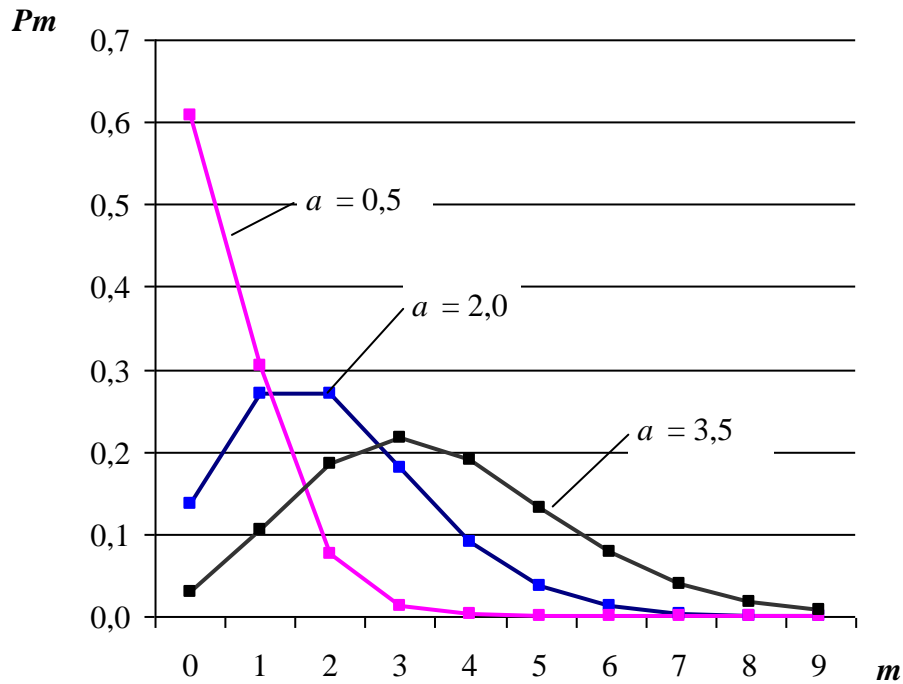


Рис. 6.5. Многоугольники распределения

Используя производную функцию дискретной СВ X , можно показать, что дисперсия тоже равна a , т.е. $D_x = a$. Таким образом, дисперсия случайной величины, распределенной по закону Пуассона, равна ее математическому ожиданию a :

$$m_x = D_x = a. \quad (6.7)$$

Это свойство распределения Пуассона часто применяется на практике для решения вопроса, правдоподобна ли гипотеза о том, что случайная величина X распределена по закону Пуассона. Для этого определяют выборочные математическое ожидание и дисперсию. Если их значения близки, то это может служить основанием в пользу гипотезы о пуассоновском распределении.

В задачах по вычислительной технике наибольший интерес представляют потоки событий, распределение которых описывается законом Пуассона.

Под *потоком событий* понимается последовательность событий, происходящих одно за другим в какие-то моменты времени. Примерами могут служить: поток сбоев электронной вычислительной машины; поток включений приборов в бытовой электросети; поток вызовов на телефонной станции; потоки отказов энергетических объектов (выключателей, разъединителей, трансформаторов и др.) в достаточно большой системе и т. п.

События, образующие поток, в общем случае могут быть различными, но мы будем рассматривать лишь поток *однородных событий*, различающихся только моментами появления. Такой поток можно изобразить как последовательность точек t_1, t_2, \dots, t_k на числовой оси (рис. 6.6).

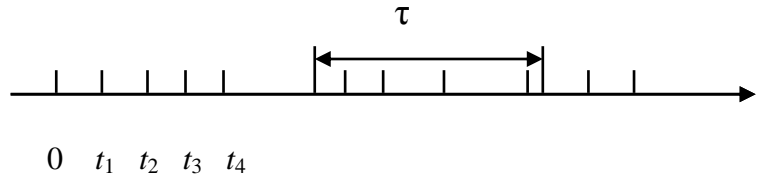


Рис. 6.6. Представление потока событий

Рассмотрим на оси $0t$ простейший поток событий как неограниченную последовательность случайных точек (рис. 6.6). Выделим произвольный участок времени длиной τ . При некоторых условиях (стационарности, отсутствия последействия и ординарности потока событий) число точек, попадающих на участок τ , распределено по закону Пуассона с математическим ожиданием:

$$a = \lambda\tau, \quad (6.8)$$

где λ – плотность потока (среднее число событий, приходящееся на единицу времени).

Вероятность того, что за время τ произойдет ровно m событий, будет равна

$$P_m(\tau) = \frac{(\lambda\tau)^m}{m!} e^{-\lambda\tau}. \quad (6.9)$$

Вероятность того, что участок окажется пустым (не произойдет ни одного события), равны

$$P_0(\tau) = e^{-\lambda\tau}. \quad (6.10)$$

Вероятность появления хотя бы одного события:

$$P_{m \geq 1}(\tau) = P_1 + P_2 + \dots + P_\infty = \sum_{m=1}^{\infty} P_m = 1 - P_{(m=0)}. \quad (6.11)$$

Вероятность того, что в интервале времени τ произойдет не менее k событий:

$$P_{m \geq k}(\tau) = P_k + P_{k+1} + P_{k+2} + \dots + P_\infty = \sum_{m=k}^{\infty} P_m = 1 - \sum_{m=0}^{k-1} P_m. \quad (6.12)$$

Распределение Пуассона является однопараметрическим. Важным свойством закона Пуассона является его предельность для биномиального распределения $X_{n,p}$. Если одновременно устремлять число опытов n к бесконечности, а вероятность p – к нулю, причем их произведение np сохраняет постоянное значение:

$$np = a. \quad (6.13)$$

Это предельное свойство биномиального распределения можно записать в виде

$$\lim_{n \rightarrow \infty, p \rightarrow 0} C_n^m p^m (1-p)^{n-m} = \frac{a^m}{m!} e^{-a}. \quad (6.14)$$

Предельное свойство биномиального закона часто находит применение на практике. Допустим, что производится большое количество n независимых опытов, в каждом из которых событие A имеет очень малую вероятность p .

Тогда для вычисления вероятности P_n^m того, что событие A появится равно m раз, можно воспользоваться приближенной формулой:

$$P_n^m \approx \frac{(np)^m}{m!} e^{-np}, \quad (6.16)$$

где $np = a$ – параметр закона Пуассона, которым приближенно заменяется биномиальное распределение.

От этого свойства закона Пуассона – выражать биномиальное распределение при большом числе опытов и малой вероятности события – происходит его название: *закон редких явлений*. Биномиальное распределение сводится в пределе к двум – нормальному и пуассоновскому.

Табличные значения функции распределения Пуассона можно получить с помощью функций *Microsoft Excel* Пуассон($m; a$; интегральное). Если аргумент «интегральное» имеет значение «истина», то функция Пуассон возвращает значение интегрального распределения $P(X \leq m) = \sum_{i=0}^m P_i$. Если аргумент «интегральное» имеет значение «ложь», то функция возвращает $P(X = m) = P_m$.

6.1.6. Распределение Эрланга

Распределение Эрланга – это гамма-распределение при целочисленном значении параметра формы k . Распределению Эрланга подчиняются суммы квадратов модулей независимых комплексных гауссовских случайных величин с нулевыми средними значениями и одинаковыми дисперсиями, поэтому распределение Эрланга часто встречается в теории надежности, инженерных приложениях, особенно телефонии.

Распределение Эрланга k -го порядка является универсальным распределением. Путем изменения целочисленного параметра k можно трансформировать данное распределение случайной величины в одно из известных. Например, при $k=1$ это распределение трансформируется в экспоненциальное, при $k=2$ – в распределение Релея, при $k > 10$ – в нормальное распределение. Выбор того или иного распределения, а следовательно, задание параметра k , определяется результатами обработки статистических данных.

Основные характеристики гамма-распределения

Область определения	$0 \leq x < \infty$
Параметры	k – параметр формы; θ – параметр масштаба
Плотность $f(x; k; \theta)$ при целых k	$\frac{1}{\theta(k-1)!} \left(\frac{x}{\theta}\right)^{k-1} e^{-\frac{x}{\theta}}$
Математическое ожидание	$k\theta$
Дисперсия	$k\theta^2$

Табличные значения функции распределения Эрланга можно получить с помощью функций *Microsoft Excel* ГаммаРасп($x; k; \theta$; интегральная). Если

значение аргумента «интегральная» – «ложь», то получаем функцию плотности $f(x; k; \theta)$.

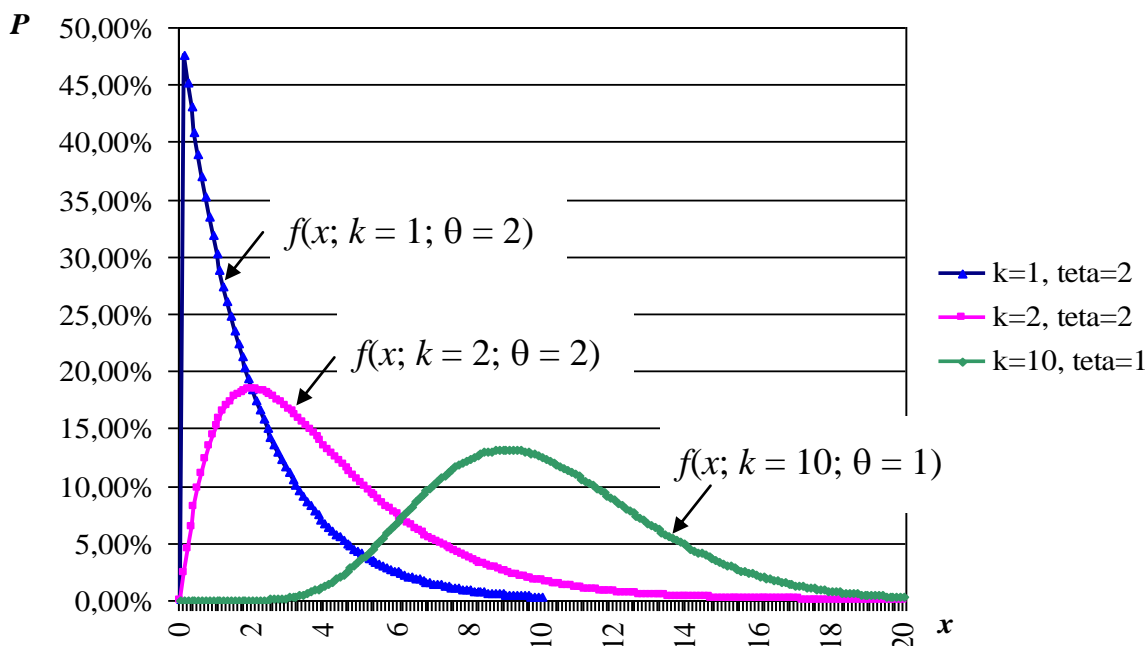


Рис. 6.7. Распределение Эрланга

6.2. ВЫБОРОЧНЫЕ РАСПРЕДЕЛЕНИЯ

После проведения экспериментов в распоряжении исследователя оказывается определенное ограниченное количество результатов измерений, характеризующих объект исследования по интересующему его признаку. Из их анализа должен быть сделан вывод, относящийся ко всем объектам данного вида. Множество объектов, конечное или бесконечное, относительно которого делаются статистические выводы, носит название генеральной совокупности. Этот термин приобретает смысл в сочетании с понятием о выборке, т. е. части этого множества. Реально же мы имеем дело с конечными генеральными совокупностями, размеры которых могут колебаться в очень широких пределах.

Например, вся генеральная совокупность уссурийских тигров насчитывает около двухсот особей, в то время как число таких деревьев, как березы или осины, измеряется миллионами. В статистических выводах важен не сам по себе объем генеральной совокупности, а та доля от него, которую составляет выборка. Во всех случаях, когда объем выборки меньше сотой части всей генеральной совокупности, последнюю по отношению к выборке принято считать практически бесконечной и использовать математический аппарат, основанный на таком представлении.

Ряд показателей, которые уже были рассмотрены: среднее, дисперсия и т. д., характеризуют выборку и носят название статистик. Такие же показатели, но характеризующие всю генеральную совокупность в целом, носят название

параметров. Таким образом, можно считать, что статистики служат оценками для параметров. Чтобы отличать статистики от параметров, принято последние обозначать греческими буквами: μ – среднее, σ^2 – дисперсия, ρ – коэффициент корреляции и т. д.

Пусть теперь мы извлекаем из генеральной совокупности большое число выборок, для каждой из которых вычисляется некоторая статистика, например среднее. Если для множества полученных средних (их будет столько, сколько было извлечено выборок) построить гистограмму, а затем аппроксимировать ее непрерывной кривой, выбрав масштаб так, чтобы площадь под кривой равнялась единице, то получим приближенное выборочное распределение среднего. Естественно, что такие выборочные распределения могут быть построены для любых интересующих нас статистик.

Предположим, что в очень большой генеральной совокупности признаки x_i , принимающие численные значения от единицы до десяти, встречаются одинаково часто. Тогда плотность распределения этих признаков будет иметь вид, как на рис. 6.8.

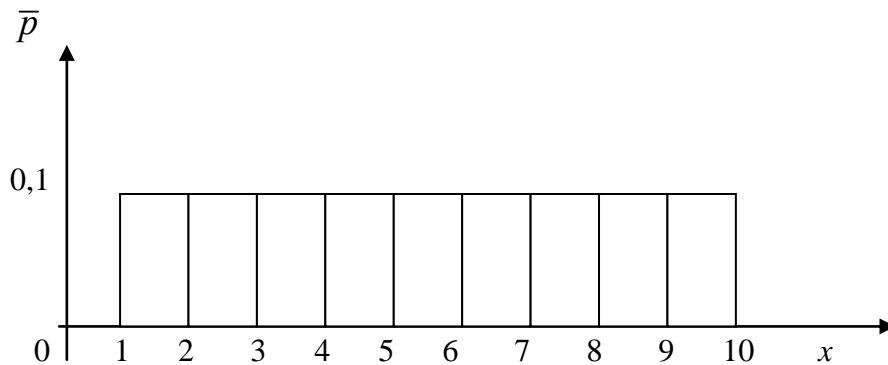


Рис. 6.8. Распределение признаков в генеральной совокупности, принимающих численные значения от единицы до десяти

Пусть теперь из всей совокупности объектов с этими признаками извлекаются 100 выборок по два объекта в каждой. Если теперь для средних этих выборок \bar{x}_i построить распределение, то оно будет иметь вид, как на рис. 6.9.

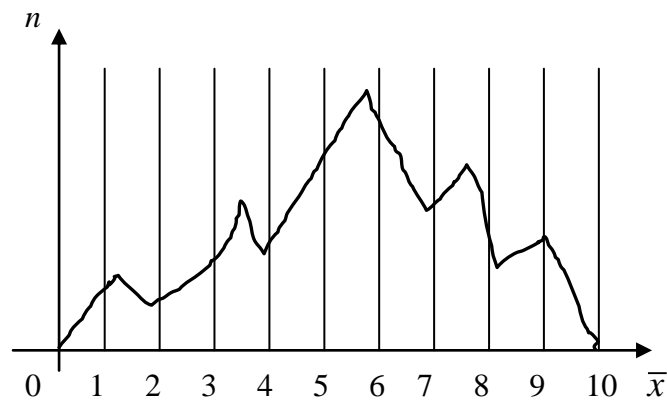


Рис. 6.9. Распределение для средних выборок

Возникает вопрос, случайно ли, что это распределение более или менее симметрично и имеет такой вид, что если его аппроксимировать непрерывной

кривой, то эта кривая будет приближаться к колоколообразной? Оказывается, нет. В специальной теореме теории вероятностей, теореме Ляпунова, утверждается, что если из бесконечной генеральной совокупности с произвольным распределением и параметрами μ и σ^2 извлекаются случайные выборки объема n , где n достаточно велико, то выборочное среднее будет иметь распределение, приближающееся к нормальному с параметрами μ и σ^2/n (σ^2 – дисперсия генеральной совокупности).

Величина σ^2/n называется дисперсией ошибки среднего, арифметическое значение корня квадратного из нее называется стандартной ошибкой среднего, т. е. $\sigma_{\bar{x}} = \sigma/\sqrt{n}$.

Рассмотрим некоторое распределение X , плотность которого представлена на рис. 6.10. Видно, что кривая подходит очень близко к оси абсцисс в точках с координатами $M[\bar{x}] \pm 3\sigma_{\bar{x}}$.

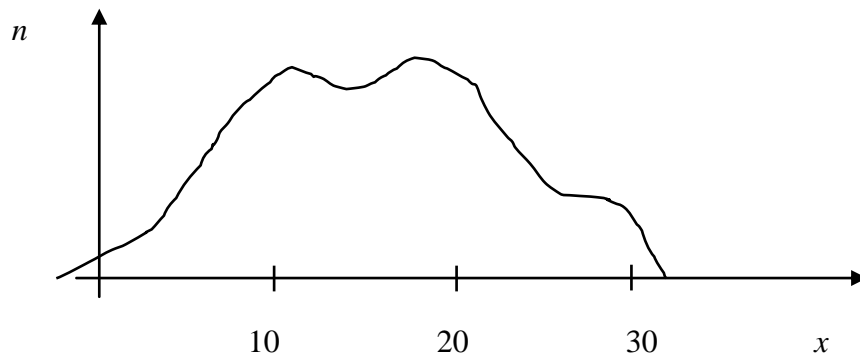


Рис. 6.10. Плотность распределения генеральной совокупности с $\mu = 15$, $\sigma_x = 10$

Построим выборочные средние \bar{x} . Когда объемы выборок равняются 100, тогда случайная величина \bar{x} имеет плотность распределения, как на рис. 6.11.

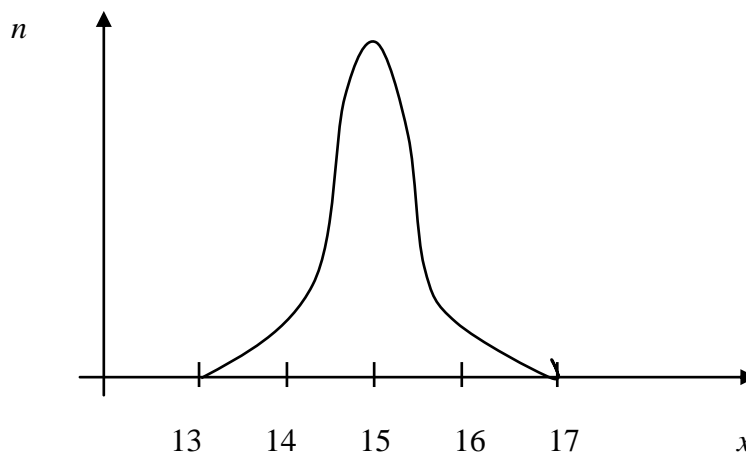


Рис. 6.11. Плотность распределения средних при объеме выборки, равном 100, $M[\bar{x}] = 15$, $\sigma_{\bar{x}}^2 = 1$

Другой пример выборочного распределения – распределение выборочных коэффициентов корреляции r , полученных из выборок двумерной нормальной генеральной совокупности с $\rho = 0$ (рис. 6.12). Мощность выборки равна 100. Тогда $M[r] = 0$, $\sigma_r = 0,1$

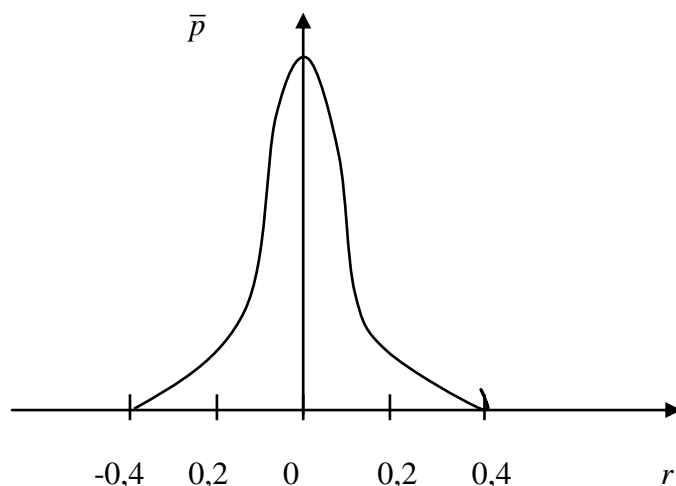


Рис. 6.12. Выборочное распределение коэффициента корреляции при $M[r] = 0$, $\sigma_r = 1$

Распределение выборочного коэффициента корреляции в этом случае стремится к нормальному распределению с нулевым средним и стандартной ошибкой $\sigma_r = 1/\sqrt{n-1}$. Оценим величину отклонения выборочных коэффициентов корреляции от нуля. Возьмем 5 % наихудших выборочных коэффициентов корреляций, т.е. больше всего отклоняющихся от теоретического $\rho = 0$. Найдем величину отклонения этих коэффициентов. Составим в *MS Excel* уравнение $\text{НОРМРАСП}(x; 0; 0,1; 1) = 1 - 0,05/2$. Здесь четвертый параметр равен 1 – в левой части стоит функция распределения. Решая уравнение, получаем величину отклонения, равную 0,22.

6.3. ИДЕНТИФИКАЦИЯ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Допустим, имеется выборочная совокупность данных $X_{\text{выб}} = \{x_1, \dots, x_n\}$ и параметрическое семейство одномерных СВ $F(p_1, \dots, p_m)$, заданных плотностями распределений $f(x, p)$, $p \in R^m$. Параметры p связаны между собой соотношениями

$$\varphi_i(p) = 0, i = \overline{1, k}. \quad (6.17)$$

Таким образом, F является m -параметрическим семейством с $m - k$ степенями свободы.

Необходимо подобрать оптимальный набор $p^{\text{опт}} \in R^m$, наилучшим образом приближающий выборку $X_{\text{выб}}$.

6.3.1. Использование метода наименьших квадратов

Построим множество интервалов $\{k_1, \dots, k_s\}$, покрывающих множество $X_{\text{выб}}$. Интервал представляет собой полузамкнутый отрезок без правой точки, они не пересекаются, их объединение также является интервалом. Длины интервалов, как правило, одинаковы и их количество рекомендуется выбирать по правилу Стерджеса

$$s = 1 + 3,33 \lg n.$$

Построим выборочную функцию плотности распределения $f_{\text{выб}}$, определенную на серединах y_j интервалов или на их концах. Значение $f_{\text{выб}}(y_j)$ равно доле точек из $X_{\text{выб}}$, попавших в соответствующий интервал k_j , $j = \overline{1, s}$. Сумма всех интервалов k_j равна L .

Рассмотрим функцию невязки

$$\Psi(p) = \sum_{j=1}^s |f_{\text{выб}}(y_j) - f_{\text{точ}}(y_j, p)|^2.$$

Решая задачу $\min \Psi(p)$ при условиях (6.17), находим $p^{\text{опт}}$. Функция Ψ может иметь несколько точек локального минимума.

6.3.2. Использование числовых характеристик выборки

Основными числовыми характеристиками выборки являются $M_{\text{выб}}$, $D_{\text{выб}}$, $\sigma_{\text{выб}}$, коэффициент скоса $A_{\text{выб}}$ и эксцесс $E_{\text{выб}}$. Аналогичные теоретические характеристики распределений $F(p)$ определяются моментами 1 – 4 порядков. Для их вычисления необходимо знать функциональное представление плотности распределения. Например,

$$A_{\text{теор}}(p) = \frac{1}{\sigma^3(p)} \int_{-\infty}^{+\infty} (x - M(p))^3 f(x, p) dx, \quad p \in R^m.$$

Интегрирование можно произвести в *MS Excel*. Для этого необходимо написать на языке *Visual Basic for Application (VBA)* функцию для вычисления известной плотности $f(x, p)$. Параметры p содержатся в диапазоне ячеек листа, а само интегрирование производится численно на отрезке, содержащем область существенных значений функции $f(x, p)$. Этот отрезок в свою очередь зависит от p .

Решим систему нелинейных уравнений

$$\begin{cases} M_{\text{теор}}(p) = M_{\text{выб}}, \\ \sigma_{\text{теор}}(p) = \sigma_{\text{выб}}, \\ A_{\text{теор}}(p) = A_{\text{выб}}, \\ E_{\text{теор}}(p) = E_{\text{выб}}, \\ \varphi_i(p) = 0, \quad i = \overline{1, k}. \end{cases} \quad (6.18)$$

Для того чтобы система (6.18) не оказалась избыточной, следует взять столько числовых характеристик, какова степень свободы $m - k$ семейства $F(p)$.

Если система (6.18) несовместна, то можно решить задачу минимизации невязки

$$\min \{ (M_{\text{теор}}(p) - M_{\text{выб}})^2 + (\sigma_{\text{теор}}(p) - \sigma_{\text{выб}})^2 + (A_{\text{теор}}(p) - A_{\text{выб}})^2 + (E_{\text{теор}}(p) - E_{\text{выб}})^2 \}$$

при условии $\varphi_i(p) = 0, \quad i = \overline{1, k}$.

ПРИМЕР 6.1. Построим $X_{\text{выб}}$ из 400 реализаций треугольного распределения $X = 1 + 2 \cdot R + 2 \cdot R$, где R – равномерно распределена на интервале $[0, 1]$.

Семейство треугольных распределений $F(p)$ зависит от $m = 4$ параметров a, b, c, d , связанных одним соотношением ($k = 1$), которое обеспечивает площадь под фигурой, равную 1 (рис. 6.13). Число степеней свободы этого семейства равно $m - k = 3$. В общем случае плотность $f(x, p)$ может быть несимметрична.

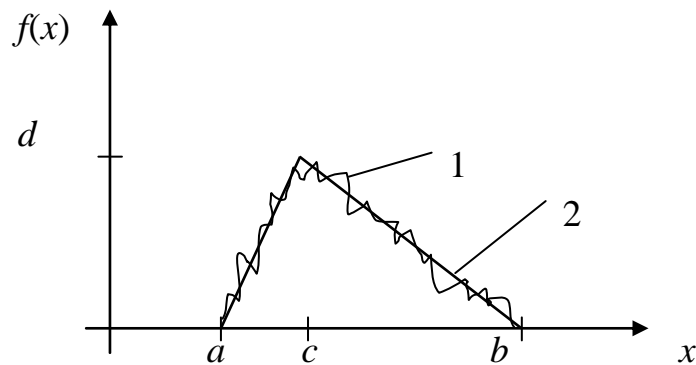


Рис. 6.13. Плотность треугольного распределения:
1 – выборочная плотность; 2 – теоретическая плотность приближенного распределения

Точные значения параметров теоретического распределения равны $a = 1, b = 5, c = 3, d = 0,5$. Найдем эти параметры двумя способами.

1. Метод наименьших квадратов стартует со значений $a = 0, b = 3, c = 1, d = 1$ (рис. 6.14, 1). Размер карманов 0,2. Количество карманов 32 от 0 до 6,2. Суммарная квадратичная невязка равна 5,33 (рис. 6.14).

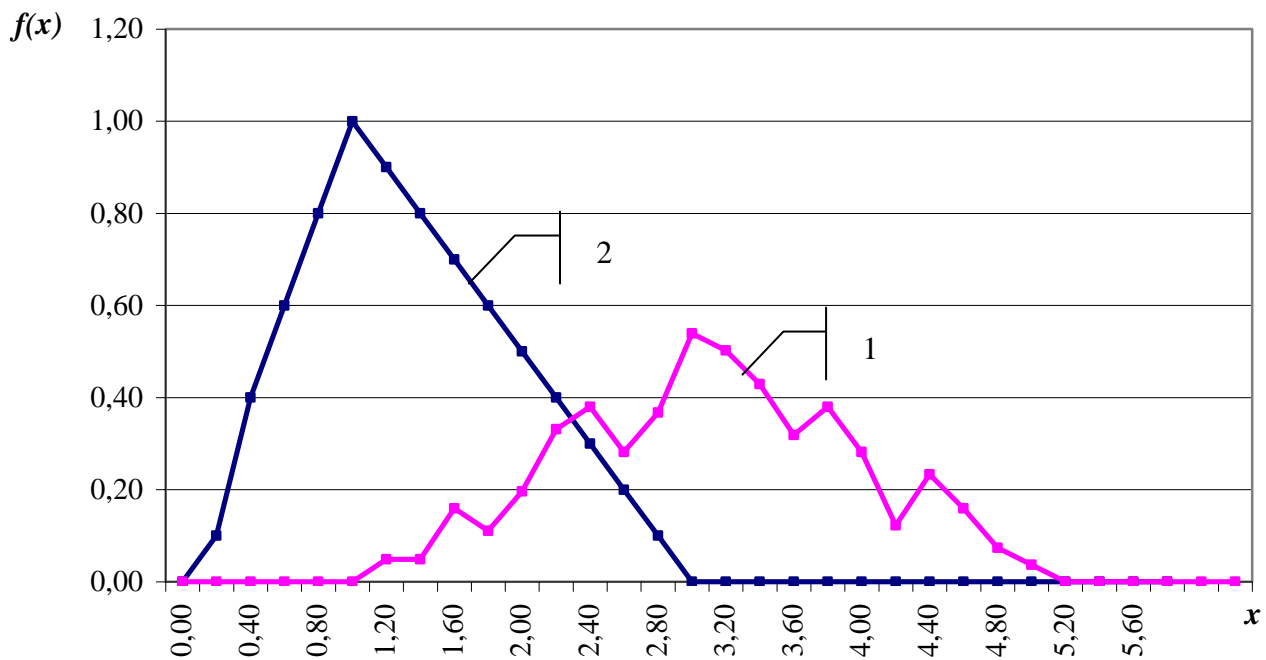


Рис. 6.14. Плотности распределений:
1 – выборочная; 2 – начальная теоретическая

В результате расчетов получили $a = 1,11, b = 5,14, c = 3,11, d = 0,50$. Суммарная квадратичная невязка равна 0,05 (рис. 6.15).

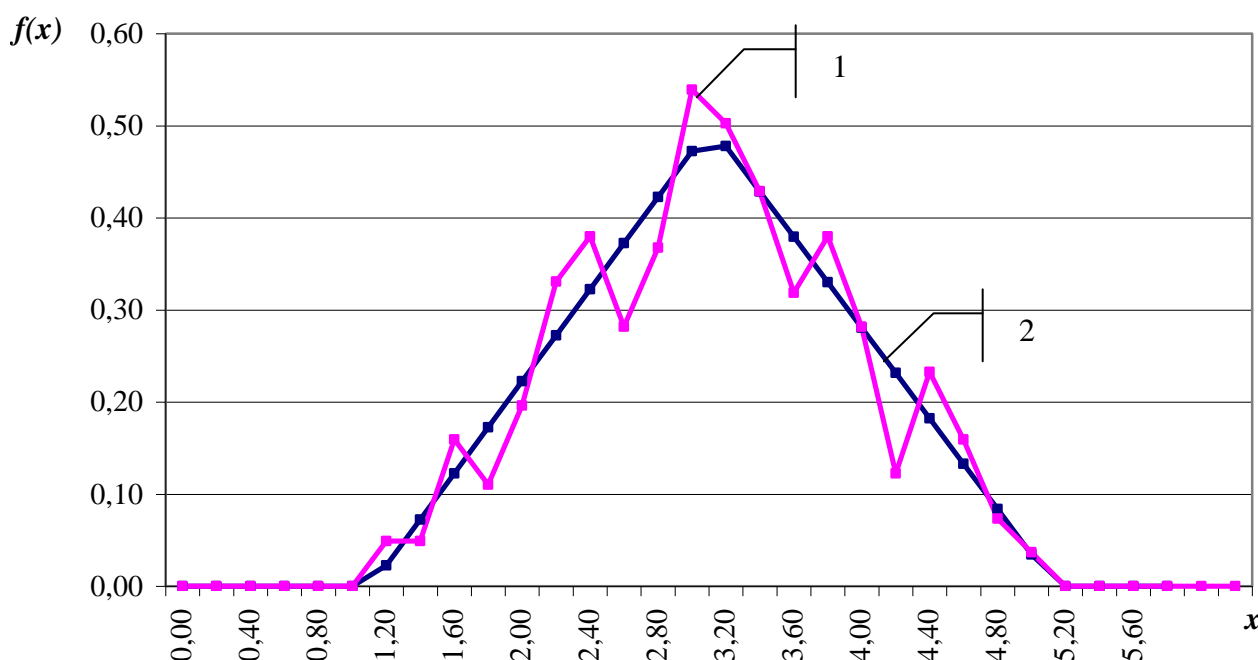


Рис. 6.15. Использование МНК:
1 – выборочная плотность; 2 – приближенная теоретическая

2. Воспользуемся числовыми характеристиками выборки.

Плотность треугольного распределения имеет вид

$$f(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{c-a}d, & a < x < c, \\ \frac{x-b}{c-b}d, & c < x < b, \\ 0, & x > b. \end{cases}$$

$$\varphi(a, b, c, d) = \frac{1}{2}(b-a)d - 1 = 0.$$

В результате расчетов выборочная плотность имеет вид графика 1 на рис. 6.16.

В систему уравнений (6.18) включим характеристики M , σ , A . Параметры распределения являются ячейками листа. Теоретические значения этих характеристик вычисляются путем численного интегрирования, а выборочные значения – с помощью функций *Microsoft Excel* СРЗНАЧ, СТАНДОТКЛ, СКОС.

Решение системы (6.18) осуществляется сервисной функцией *Microsoft Excel* Сервис/Поиск решения.

Начальные значения параметров распределения

$$a = 0, b = 3, c = 1, d = 1.$$

Расчеты показали следующее. Найденные значения параметров распределения

$$a = 0,79, b = 5,26, c = 3,01, d = 0,45.$$

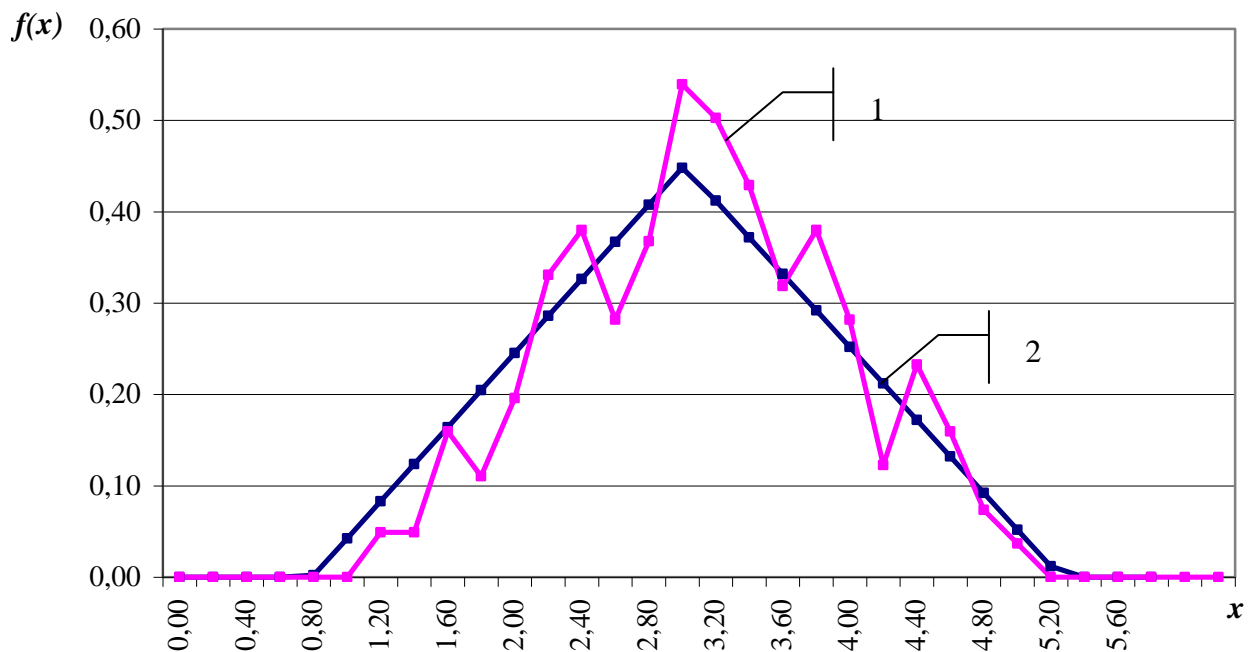


Рис. 6.16. Использование числовых характеристик:
1 – выборочная плотность; 2 – приближенная теоретическая плотность

Выборочные характеристики равны

$$M_{\text{выб}} = 3,02; \sigma_{\text{выб}} = 0,83; A_{\text{выб}} = 0,01; E_{\text{выб}} = -0,51.$$

Теоретические приближенные характеристики равны

$$M_{\text{прибл}} = 3,02; \sigma_{\text{прибл}} = 0,83; A_{\text{прибл}} = 0,01; E_{\text{прибл}} = 0,84.$$

Точные теоретические характеристики рассчитаны с помощью вероятностного калькулятора (см. гл. 12).

$$M_{\text{теор}} = 3,00; \sigma_{\text{теор}} = \sqrt{\frac{2}{3}} = 0,82; A_{\text{теор}} = 0; E_{\text{теор}} = -0,60.$$

6.4. ГЕНЕРАЦИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

6.4.1. Метод «обратной» функции

Пусть имеется одномерная непрерывная СВ X с функцией распределения $y = F(x)$. Необходимо построить алгоритм генерации этой СВ.

Возьмем одномерную СВ Y , равномерно распределенную на $[0; 1]$. Это распределение реализовано во всех программных пакетах. Например, на языке C++

$$Y = \text{random}(1001)/1001.0.$$

На языке VBA:

$$Y = \text{слчис}().$$

Алгоритм реализуется в два этапа.

Этап 1. Получить реализацию \tilde{y} величины Y .

Этап 2. Решить уравнение $F(x) = \tilde{y}$ относительно x . Решение $\tilde{x} = F^{-1}(\tilde{y})$

– искомая реализация X .

6.4.2. Генерация двумерных случайных величин

Пусть СВ $Z = (X, Y)$ задана двумерной функцией распределения $z = F(x, y)$, которая задается формулой $F(x, y) = P(X < x \text{ и } Y < y)$.

Свойства двумерной функции распределения $F(x, y)$:

1) ограниченность:

$$0 \leq F(x, y) \leq 1;$$

2) предельность:

$$F(-\infty, -\infty) = F(-\infty, y) = F(x, -\infty) = 0, F(+\infty, +\infty) = 1;$$

3) вероятность попадания в диапазон:

$$P(a_1 \leq X < a_2 \text{ и } b_1 \leq Y < b_2) = F(a_2, b_2) - F(a_1, b_2) - F(a_2, b_1) + F(a_1, b_1).$$

Определение 1. Предельными распределениями для Z называются одномерные распределения $F(x, -\infty)$ и $F(-\infty, y)$. Предельные распределения задают две одномерные СВ, для которых можно применять метод «обратных» функций.

Определение 2. Условными распределениями называются одномерные распределения

$$F(x|Y = \tilde{y}) = \frac{F(x, \tilde{y})}{F(+\infty, \tilde{y})}, \quad (6.19)$$

$$F(y|X = \tilde{x}) = \frac{F(\tilde{x}, y)}{F(\tilde{x}, +\infty)}. \quad (6.19a)$$

Знаменатели в формулах (6.19), (6.19a) играют нормирующую роль и обеспечивают предельное свойство функции распределения $F(+\infty | Y = \tilde{y}) = 1$.

Алгоритм генерации величины Z

Этап 1. Пусть \tilde{r}_1 – реализация одномерной СВ, равномерно распределенной на $[0; 1]$.

Этап 2. Решим уравнение $F(x, +\infty) = \tilde{r}_1$ и найдем его решение \tilde{x} .

Этап 3. Найдем другую реализацию \tilde{r}_2 одномерной СВ, равномерно распределенной на $[0; 1]$.

Этап 4. Решим уравнение $F(y | X = \tilde{x}) = \tilde{r}_2$ и найдем его решение \tilde{y} .

Пара чисел (\tilde{x}, \tilde{y}) – искомая реализация двумерной СВ Z .

ПРИМЕР 6.2. Совместная плотность распределения СВ Z равна $F(x, y) = xy^2$ на множестве $[0; 1] \times [0; 1]$.

Если $x < 0$ или $y < 0$, то $F = 0$,
иначе если $x > 1$ и $y < 1$, то $F = y^2$,
иначе, если $x < 1$ и $y > 1$, $F = x$,
иначе $F = 1$.

Построим генератор СВ Z .

Найдем предельные распределения

$$F(x, +\infty) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x \leq 1, \\ 1, & x > 1. \end{cases}$$

$$F(+\infty, y) = \begin{cases} 0, y < 0, \\ y^2, 0 \leq y \leq 1, \\ 1, y > 1. \end{cases}$$

Условные распределения

$$F(x|Y = \tilde{y}) = \begin{cases} 0, x < 0, \\ \frac{x\tilde{y}^2}{\tilde{y}^2} = x, 0 \leq x \leq 1, \\ 1, x > 1. \end{cases}$$

$$F(y|X = \tilde{x}) = \begin{cases} 0, y < 0, \\ \frac{\tilde{x}y^2}{\tilde{x}} = y^2, 0 \leq y \leq 1, \\ 1, y > 1. \end{cases}$$

На этапе 2 решаем уравнение $F(x, +\infty) = \tilde{r}_1$. $x = \tilde{r}_1$. Находим решение $\tilde{x} = \tilde{r}_1$.

На этапе 4 решаем уравнение $F(y | X = \tilde{x}) = \tilde{r}_2$. $y^2 = \tilde{r}_2$. Находим решение $\tilde{y} = \sqrt{\tilde{r}_2}$.

Ответ: $(\tilde{r}_1, \sqrt{\tilde{r}_2})$ – искомая реализация СВ.

Определение 3. Плотностью распределения СВ $Z = (X, Y)$ называется функция

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}. \quad (6.20)$$

Из формулы (6.20) вытекает

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy.$$

Изложенный алгоритм можно сформулировать в терминах функции плотности $f(x, y)$.

Таким образом, продемонстрирован следующий подход к статистической обработке одномерных выборочных данных.

1. Из большого количества семейств распределений выбираем несколько наиболее подходящих (раздел 6.1).

2. Для каждого семейства находим распределение, наиболее точно приближающее выборку (разделы 6.3, 6.4).

3. Используя критерии согласия, установить совпадение выборочного и теоретического распределений [13].

4. Используем полученное теоретическое распределение для генерации новых реализаций с целью проведения имитационного моделирования (раздел 6.4.1).

6.5. ИНТЕРВАЛЬНОЕ ОЦЕНИВАНИЕ

Различные статистики, получаемые в результате вычислений, представляют собой точечные оценки соответствующих параметров. В предыдущем параграфе указывалось, что если из генеральной совокупности извлечь некоторое количество выборок и для каждой из них найти интересующую нас статистику, то вычисленные значения будут представлять собой случайные величины, имеющие некоторый разброс вокруг оцениваемого параметра. Как правило, в результате эксперимента в распоряжении исследователя имеется одна выборка, поэтому значительный интерес представляет получение интервальной оценки, т. е. интервала, внутри которого, как можно предположить, лежит истинное значение параметра.

Для примера интервального оценивания рассмотрим как \bar{x} оценку μ . Нам известно, что если выборки извлекаются из генеральной совокупности с параметрами μ и σ^2 , то распределение выборочных \bar{x} будет иметь среднее, равное μ , и дисперсию σ^2/n и будет стремиться к нормальному (рис. 6.17).

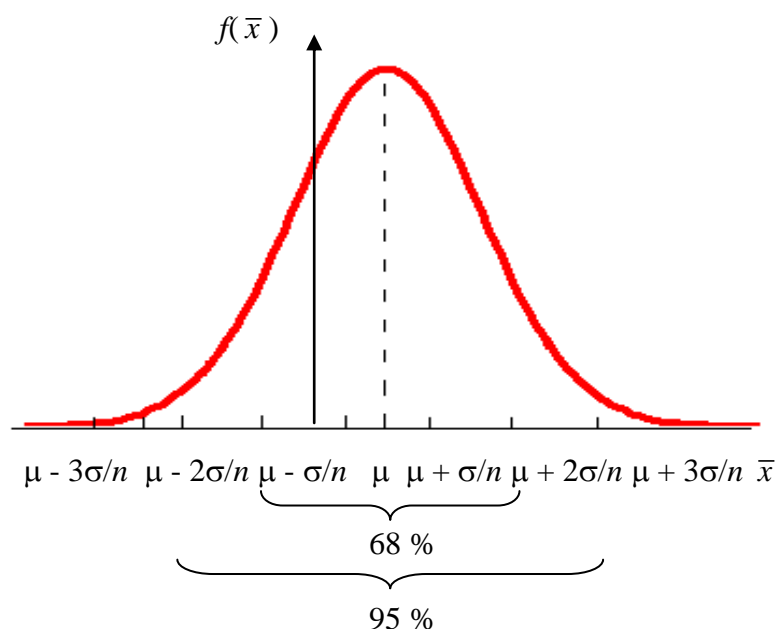


Рис. 6.17. Распределение выборочных \bar{x} , имеющее среднее, равное μ , дисперсию σ^2/n

Для такого распределения 68 % наблюдений лежит в пределах одного стандартного отклонения, относительно μ , т. е. в пределах $\mu \pm \sigma/\sqrt{n}$ (стандартное отклонение выборочного распределения \bar{x} не что иное, как стандартная ошибка среднего). Соответственно в пределах $\mu \pm 2\sigma/\sqrt{n}$ и $\mu \pm 3\sigma/\sqrt{n}$ лежит около 95 и 99 % всех возможных значений выборочных средних. (Коэффициенты 2 и 3 взяты для упрощения, точные их значения равны 1,96 и 2,58). Отсюда следует вывод, что если, например, 95 % значений \bar{x} расположено внутри интервала $\pm 2\sigma/\sqrt{n}$ относительно μ , то $\bar{x} \pm 2\sigma/\sqrt{n}$, где \bar{x} вычислено по одной выборке из генеральной совокупности с параметрами μ и σ^2 , дает интервал, который включает μ в своих границах (рис. 6.17).

Сделанный вывод допускает и следующую формулировку: вероятность P того, что расстояние от \bar{x} до μ , измеренное в единицах стандартной ошибки, больше или меньше $z_{1-\frac{\alpha}{2}}$, равна $1-\alpha$, т.е.

$$P\left(\bar{x} - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha. \quad (6.21)$$

Для выборок достаточно большого объема распределение \bar{x} приближается к нормальному. Следовательно, используя соответствующую таблицу площадей под кривой стандартного нормального распределения, можно определить значения $\pm z$, между которыми заключена интересующая нас часть площади. Оставшуюся часть площади обозначим через α . В силу того, что кривая распределения симметрична, слева от $-z$ и справа от z будет находиться по $\alpha/2$ части площади стандартного нормального распределения. Табличные значения z даны в единицах стандартного отклонения, которое для стандартизованного нормального распределения равно единице. Следовательно, чтобы найти интересующую нас часть площади $1-\alpha$ под кривой нормального распределения с произвольными параметрами, необходимо табличные значения $z_{1-\frac{\alpha}{2}}$ умножить на величину стандартного отклонения, которое в нашем случае равно σ/\sqrt{n} . Величина $1-\alpha = P$ носит название доверительной вероятности, а сам доверительный интервал, т.е. интервал, внутрь которого с заданной доверительной вероятностью попадает истинное значение параметра (рис. 6.18), задается выражением

$$\bar{x} \pm z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}. \quad (6.22)$$

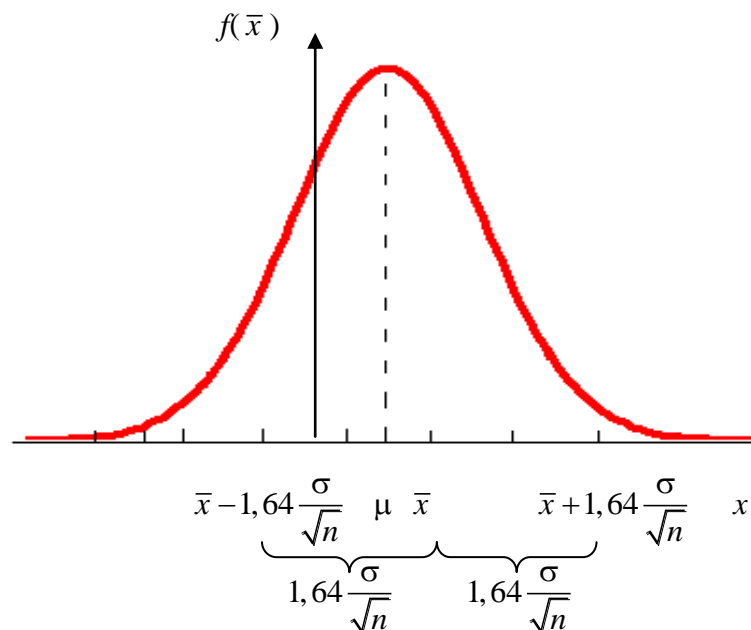


Рис. 6.18. Распределение выборочных \bar{x} и доверительный интервал $\bar{x} \pm 1,64 \frac{\sigma}{\sqrt{n}}$

ПРИМЕР 6.3. Пусть $n = 144$, $\sigma^2 = 49$, а доверительная вероятность произвольно выбрана равной 0,91, т. е. $P = 1 - \alpha = 0,91$. Отсюда $\alpha = 0,09$. Из [10] доля $\alpha/2 = 0,045$ (или 4,5 %) площади под единичной нормальной кривой лежит слева от значения $z = 1,7$. Следовательно, $z_{0,955} = 1,7$. Пусть среднее этой выборки равно 18,2. Тогда доверительный интервал для μ равен:

$$\bar{x} \pm z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = 18,2 \pm 1,7 \frac{7}{12} = 18,2 \pm 0,992 = (17,208; 19,192).$$

Замечание. На практике дисперсия генеральной совокупности неизвестна. Поэтому вместо σ^2 обычно используют s^2 , вычисленную по выборке, и вместо таблиц стандартного нормального распределения пользуются таблицами распределения Стьюдента.

Рассмотренные принципы построения доверительных интервалов справедливы, естественно, не только для средних, но и для всех других статистик, для которых показано, что распределение их выборочных значений приближается к нормальному.

Требование нормальности является принципиальным, так как только при его выполнении имеет смысл приведенная выше процедура построения доверительных интервалов. Как правило, для большинства статистик это требование выполняется, но не для всех. В частности, нормальность выборочных коэффициентов корреляции имеет место только для случая, когда в двумерной генеральной совокупности, из которой извлекаются выборки с последующим вычислением r_{xy} , $\rho = 0$. Для других значений ρ распределение выборочных коэффициентов корреляции имеет большую асимметрию и может быть аппроксимировано кривой нормального распределения (рис. 6.19).

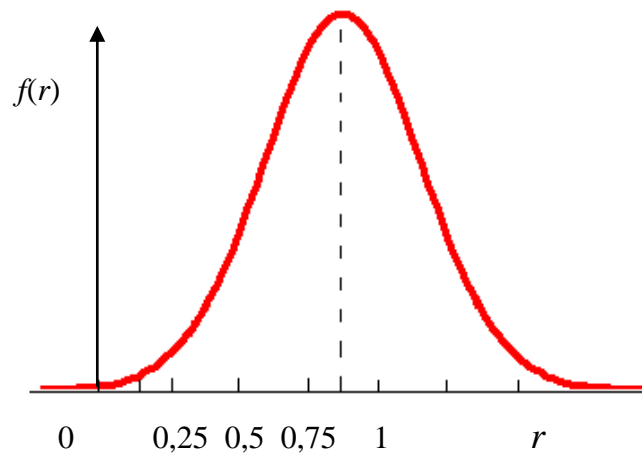


Рис. 6.19. Распределение выборочных коэффициентов корреляции для $\rho = 0,8$, $n = 20$

Тем не менее, интервальное оценивание возможно и в этом случае, но при использовании не самих выборочных коэффициентов корреляции, а некоторых преобразованных величин. Р. Фишером было показано, что величина

$$z_r = \frac{1}{2} \ln \frac{1+r}{1-r}, \quad (6.23)$$

где r – выборочный коэффициент корреляции, имеет нормальное распределение со средним, равным z_r , и стандартным отклонением $1/\sqrt{n-3}$ (здесь и в дальнейшем для простоты индекс «ху» у коэффициента корреляции опущен).

Рассмотрим пример построения доверительного интервала для r с использованием z -преобразования Фишера. Значения z_r для различных значений коэффициента корреляции приведены в [10].

Пусть число пар значений, по которым вычислялось r , равно 12, а само полученное значение коэффициента корреляции равно 0,79. Необходимо построить доверительный интервал для r , который с доверительной вероятностью $P = 0,95$ накрывал бы истинное значение параметра ρ .

Преобразуем r в z_r . Для $r = 0,79$ имеем $z_r = 1,07$. Преобразованная величина z_r , как уже было сказано выше, имеет нормальное распределение со стандартным отклонением $\sigma_z = 1/\sqrt{n-3}$. Так как $z_{1-\alpha/2} = z_{0,975} = 1,96$, то доверительный интервал для z_r будет равен:

$$z_r \pm z_{1-\alpha/2} \sigma_z = 1,07 \pm 1,96 \frac{1}{\sqrt{12-3}} = 1,07 \pm 0,653 = (0,418; 1,723).$$

Теперь, вновь используя z -преобразования (6.23), можно перейти к значениям r . Доверительные пределы для коэффициента корреляции будут приблизительно равны (0,395; 0,938).

В заключение этого параграфа приведем формулы для стандартных ошибок некоторых статистик и те условия, при выполнении которых распределения выборочных значений этих статистик приближаются к нормальным, и, следовательно, для них можно строить доверительные интервалы по описанным выше правилам.

1. Медиана (Md) – $\sigma_{Md} = \sqrt{\frac{2n}{\pi}} \sigma$. Здесь n – число значений в выборке,

σ – стандартное отклонение генеральной совокупности. Формула справедлива, если генеральная совокупность симметрична, а $n \geq 30$.

2. Стандартное отклонение (σ) – $\sigma_\sigma = \sqrt{2\pi} \sigma$. Распределение выборочных стандартных отклонений близко к нормальному, когда $n \geq 100$ и исходная генеральная совокупность нормальна.

3. Коэффициент вариации (cv) –

$$\sigma_{cv} = \frac{cv}{\sqrt{2n}} \sqrt{1 + \frac{2(cv)^2}{10^4}}.$$

Вторым сомножителем в этом выражении можно ввиду его малости пренебречь, тогда $\sigma_{cv} = \frac{cv}{\sqrt{2n}}$. Ограничения те же, что и для стандартного отклонения.

4. Коэффициенты регрессии (a и b) – $\sigma_a = \frac{\sigma}{\sqrt{n}}$, $\sigma_b = \frac{\sigma}{\sqrt{\sum (x - \bar{x})^2}}$, где

σ – стандартное отклонение зависимой переменной y , которая должна иметь обязательно нормальное распределение с параметром σ^2 для всех значений независимой переменной x .

5. Для дихотомической случайной переменной, в которой доля объектов вида A представляет собой величину $p_A = n_A/n$, где n_A и n – соответственно число объектов вида A и общее число объектов

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}},$$

где в данном случае через π обозначена доля объектов вида A в генеральной совокупности.

6.6. ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

После проведения эксперимента в распоряжении исследователя имеется некоторое множество значений измеренных показателей, и возникает вопрос, какие выводы о свойствах генеральной совокупности можно сделать по этим выборочным наблюдениям. Первым шагом в решении поставленной задачи может быть вычисление различных статистических оценок и построение для них доверительных интервалов. Вторым обязательным шагом является проверка выдвигаемых статистических гипотез [10, 14].

Нужно четко представлять себе разницу между гипотезой в обычном понимании и статистической гипотезой. *Статистическая гипотеза* – это любое предположение относительно распределения наблюдаемых случайных величин, в то время как в других областях знаний гипотеза – предположительное суждение о закономерной, причинной связи явлений.

Пусть, например, при изучении взаимосвязи между двумя переменными получена выборка из двумерной генеральной совокупности и рассчитан коэффициент корреляции. Будем считать, что эта выборка достаточно представительна (репрезентативна), т. е. хорошо отражает свойства всей генеральной совокупности. Имея в своем распоряжении выборочный коэффициент корреляции r , исследователь выдвигает статистическую гипотезу о том, что коэффициент корреляции ρ генеральной совокупности равен какому-то значению, например нулю. Формально это записывается так: $H: \rho = 0$. (Здесь и в дальнейшем буквой H обозначается статистическая гипотеза.) Процедура, которая будет использоваться для принятия решения об истинности или ложности сформулированной статистической гипотезы, называется проверкой гипотезы. Итак, если вычисленное значение $r \neq 0$, значит ли это, что H ошибочна?

Выборочные значения различных статистик, в том числе и коэффициента корреляции, представляют собой случайные величины, имеющие порой довольно большой разброс, поэтому даже в том случае, когда $\rho = 0$, вычисленные по выборке значения могут существенно отличаться от нуля.

Следовательно, утверждать с уверенностью по полученному значению r , равен или не равен нулю коэффициент корреляции ρ (и это зависит от объема выборки), экспериментатор не может. Эти соображения составляют один из основных принципов, лежащих в основе проверки статистических гипотез: при проверке любой статистической гипотезы решение никогда не принимается с абсолютной уверенностью, всегда существует риск принятия неправильного решения. Именно в контроле и оценке этого риска состоит сущность проверки статистических гипотез.

Если выборочные значения представляют собой случайные величины, то возможно получение любого значения в нашем случае для коэффициента корреляции в диапазоне ± 1 . Однако некоторые значения более, а другие менее вероятны. В частности, выборочная плотность распределения коэффициентов корреляции для выборок объемом 100 (рис. 6.20) показывает, что, если в генеральной совокупности $\rho = 0,14$, получение выборочных значений, например $r = 0,4$ или $r = -0,2$, весьма маловероятно, хотя полностью их исключить нельзя.

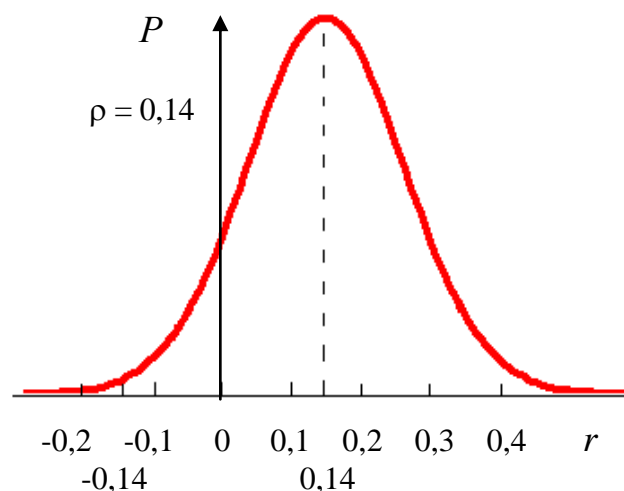


Рис. 6.20. Выборочная плотность распределения коэффициента корреляции для $\rho = 0,14$

Какой вывод можно из этого сделать? Пусть, например, получено выборочное значение $r = 0,5$. С одной стороны, такое значение должно способствовать отклонению гипотезы $H: \rho = 0$ ввиду ничтожной вероятности его появления, если истинное значение коэффициента корреляции, в самом деле, равно 0,1. Эта вероятность, конечно же, очень мала, но все-таки не равна нулю. С другой стороны, истинное значение ρ в генеральной совокупности на самом деле нам неизвестно. А если, например, $\rho = 0,3$? В этом случае появление выборочного значения $r = 0,5$ не такое уж маловероятное событие. А какое-то решение принимать надо. Поэтому вводят следующие понятия ошибки первого и второго рода.

Вернемся к примеру. Если, в самом деле, $\rho = 0$ и выдвинута гипотеза $H: \rho = 0$, но получив значение $r = 0,5$, мы ее отвергаем (а сделать это можно, поскольку и такое значение, хотя и очень редко, может появиться), возникает ошибка первого рода.

Пусть выдвинута та же гипотеза $H: \rho = 0$, в то время как истинное значение генеральной совокупности равно, например, $-0,3$, и мы, получив значение $r = 0,12$, примем ее. Появляется ошибка второго рода.

Таким образом, при проверке статистических гипотез существует четыре возможности:

1. H верна, и она принимается.
2. H верна, но она отвергается (ошибка первого рода).
3. H неверна, и она отвергается.
4. H неверна, но она принимается (ошибка второго рода).

Ошибки первого и второго рода существенно различаются между собой по значимости, и это оказывает большое влияние на всю процедуру проверки статистических гипотез.

Необходимо еще раз подчеркнуть, что никакая гипотеза не может быть окончательно принята или отвергнута. Поэтому используемые в дальнейшем категорические утверждения «принять» и «отвергнуть» являются просто условными сокращениями выражений вида «опытные данные не противоречат выдвинутой гипотезе» и «опытные данные противоречат выдвинутой гипотезе».

Основные этапы проверки статистических гипотез:

1. Формулируется проверяемая гипотеза, например $H: \rho = 0$. В силу исторической традиции проверяемая гипотеза носит название *нуль-гипотезы*.

2. Делаются некоторые предположения относительно выборочного распределения статистики, с помощью которой планируется оценивать параметр, входящий в гипотезу.

3. Принимается степень риска отвергнуть выдвинутую гипотезу, если она верна. Этот риск α называется *уровнем значимости* проверки гипотезы и выражается как вероятность. Из определения уровня значимости следует, что его величина определяет уровень ошибки первого рода. Уровень значимости определяет также *критические области*, т.е. области, попадание в которые выборочного значения статистики, оценивающей параметр, приводит к тому, что сформулированная гипотеза отвергается. Критические области и область принятия гипотезы для $H: \rho = 0$ и $\alpha = 0,05$ – приведены на рис. 6.21.

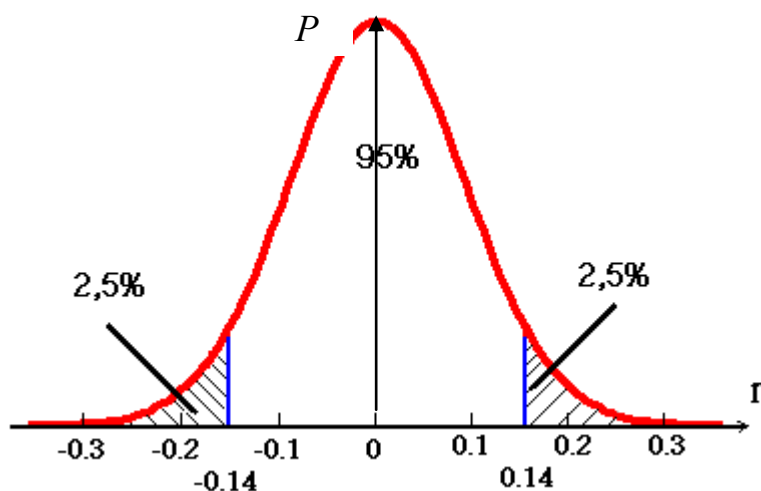


Рис. 6.21. Область принятия гипотезы (95 %) и критическая область (0,5 %)

4. Извлекается выборка; рассчитывается значение интересующей нас статистики; определяется, в какую область, критическую или допустимую, она попала, и на основании этого принимается решение относительно истинности гипотезы H . Решение принимается с использованием критерия для проверки статистической гипотезы, который представляет собой правило для определения ложности или истинности выдвинутой гипотезы.

Уровнем значимости, а значит, и вероятностью ошибки первого рода можно управлять. В принципе мы можем установить любую приемлемую для нас степень риска для неправильного вывода на основе выборочных данных об ошибочности выдвинутой гипотезы. Поэтому на первый взгляд кажется целесообразным выбирать уровень значимости как можно меньшим, так как в этом случае вероятность отвергнуть правильную гипотезу будет минимальной. Однако, сформулировав проверяемую гипотезу, мы не знаем, верна ли она. Поэтому уменьшение вероятности ошибки первого рода автоматически будет приводить к увеличению вероятности ошибки второго рода. Поясним это на примере.

Обозначим нуль-гипотезу о том, что $\rho = 0$, через H_0 : $\rho = 0$, а альтернативную ей гипотезу – H_1 : $\rho \neq 0$. Предположим, что на самом деле истинное значение $\rho = 0,20$ и извлекается выборка объемом $n = 200$. На рис. 6.22 кривые представляют собой распределения выборочных коэффициентов корреляции для генеральных совокупностей с $\rho = 0$ и $\rho = 0,2$.

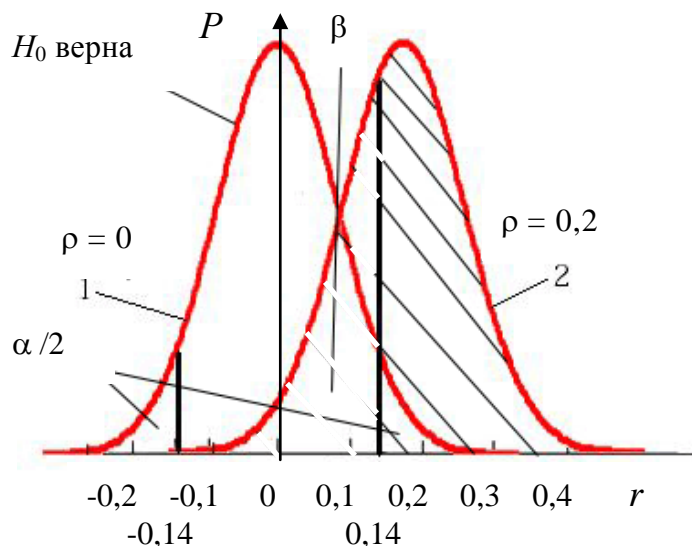


Рис. 6.22. Распределения выборочных коэффициентов корреляции для генеральных совокупностей при $\rho = 0$ (1) и $\rho = 0,2$ (2)

Из рис. 6.22 видно, что если выбрать уровень значимости $\alpha = 0,05$, то в случае истинности $\rho = 0$ появление выборочных значений, больших 0,14 или меньших $-0,14$, возможно только в пяти случаях из ста, и для таких значений гипотеза H_0 будет отвергаться. Все же остальные значения r , попадающие в интервал $(-0,14; 0,14)$, будут приводить к принятию H_0 . Но что будет, если на самом деле $\rho = 0,20$? Тогда нуль-гипотезу H_0 : $\rho = 0$ следовало бы отклонить в

пользу альтернативной гипотезы $H_1: \rho \neq 0$, и это будет делаться для значений $r > 0,14$ (область со штриховкой).

Площадь этой заштрихованной области численно равна вероятности отвергнуть неправильную гипотезу, а сама эта вероятность $(1 - \beta)$ носит название *мощности выбранного критерия*, в данном случае критерия $\rho = 0,20$.

С другой стороны, если мы приняли $H_0: \rho = 0$, в то время как справедлива $H_1: \rho = 0,20$, такие значения, как $r = 0,1$, будут свидетельствовать в пользу этой неправильной гипотезы, т.е. будет совершаться ошибка второго рода. Из рис. 6.22 видно, что площадь под кривой 2, соответствующая выборочным значениям $r < 0,14$ и перекрывающаяся с областью принятия гипотезы $H_0: \rho = 0$, довольно значительна (в данном случае она равна 18 %). Поэтому и вероятность принять неправильную гипотезу о том, что $\rho = 0$, численно равная этой площади, тоже велика. При этом мы видим, что если уровень значимости α выбрать еще меньше, то вероятность ошибки второго рода β еще больше возрастает, а мощность критерия $(1 - \beta)$ уменьшается.

Таким образом, становится ясно, что уровень значимости и мощность критерия связаны между собой, причем связь эта нелинейная. Поэтому произвольно по нашему усмотрению изменять уровень значимости нельзя, так как неоправданное уменьшение ошибки первого рода может привести к существенной потере мощности критерия для проверки статистической гипотезы.

Наиболее часто в статистической практике используются уровни значимости $\alpha = 0,05$ и $\alpha = 0,01$, потому что они обеспечивают разумный компромисс между ошибками первого и второго рода. Но абсолютизировать эти уровни значимости нельзя. Вообще, приемлемый уровень значимости необходимо выбирать, исходя из условий решаемой исследовательской задачи.

ПРИМЕР 6.4. Пусть испытывается вновь созданный лекарственный препарат, действие которого направлено на поддержание некоторого существенного для жизни показателя в определенных пределах. Испытывая этот препарат, можно, например, выдвинуть гипотезу, что он в самом деле удерживает среднее значение показателя в нужных пределах. И если был выбран уровень значимости $\alpha = 0,05$, то в случае справедливости выдвинутой гипотезы она будет отвергаться в одном случае из двадцати. Но если гипотеза неверна, то при таком уровне ошибки первого рода уровень ошибки второго рода будет достаточно высок, а это значит, что мы можем принять неправильную гипотезу во многих случаях, когда препарат не выполняет своих функций. Естественно, что это недопустимо, когда речь идет о здоровье или жизни людей. Поэтому в этих случаях нужно выбирать уровень значимости, по крайней мере, $\alpha = 0,10$ или даже еще больше, чтобы свести риск принятия неправильной гипотезы к минимальному значению. Правда, при этом мы значительно чаще будем отвергать правильную гипотезу, и возможно, что для ее окончательной проверки понадобится провести значительно больше экспериментов. Это будет, конечно, связано с дополнительными затратами, но

в случаях, аналогичных вышеприведенному, эти соображения всегда должны отодвигаться на второй план.

ПРИМЕР 6.5. С другой стороны, может встретиться прямо противоположная ситуация, когда уровень значимости нужно выбирать меньше даже чем 0,01. Это может потребоваться в тех случаях, когда проверяются статистические гипотезы о некоторых параметрах, выборочные статистики для которых получаются при проведении экспериментов, направленных на поиск новых эффектов, предсказываемых теорией или связанных с использованием уникального оборудования. В этих случаях экспериментатору очень важно не упустить эффект, если он есть. Поэтому он может предусмотреть, чтобы выдвигаемая им, и как он считает, правильная гипотеза отвергалась не более чем в одном случае из тысячи. При этом экспериментатор должен отдавать себе полный отчет в том, что существенно возрастает ошибка второго рода, т.е. вероятность принять выдвигаемую гипотезу, даже если она неверна, со всеми вытекающими отсюда последствиями.

Варьируя уровень значимости α и число наблюдений n , можно в каждом конкретном случае выбрать разумный компромисс между α и мощностью критерия $1 - \beta$, помня при этом, что мощность критерия для проверки статистической гипотезы возрастает с увеличением n и α .

7. ПРИМЕНЕНИЕ КРИТЕРИЕВ СОГЛАСИЯ

В практике статистического анализа экспериментальных данных основной интерес являет не само по себе вычисление тех или иных статистик, а ответы на следующие вопросы. Действительно ли среднее генеральной совокупности равно некоторому числу? Значимо ли отличается от нуля коэффициент корреляции? Равны ли дисперсии двух выборок? И таких вопросов в зависимости от конкретной исследовательской задачи может возникать много. Соответственно разработано и множество критериев для проверки выдвигаемых статистических гипотез. Рассмотрим некоторые наиболее употребительные из них [10]. В основном они будут относиться к средним, дисперсиям, коэффициентам корреляции и оценке резко выделяющихся наблюдений.

Все критерии для проверки статистических гипотез делятся на две большие группы: параметрические и непараметрические. Параметрические критерии основаны на предположении о том, что выборочные данные взяты из генеральной совокупности с известным распределением, и основная задача состоит в оценке параметров этого распределения. Для непараметрических критериев не требуется никаких предположений о характере распределения, за исключением предположения о том, что оно непрерывно.

Рассмотрим параметрические критерии. Последовательность проверки будет включать формулирование нуль-гипотезы H_0 и альтернативной гипотезы H_1 , формулирование сделанных допущений, определение выборочной статистики, используемой при проверке H_0 и H_1 , получение выборочного распределения проверяемой статистики, определение критических областей для

выбранного критерия и построение доверительного интервала для выборочной статистики.

7.1. КРИТЕРИИ СОГЛАСИЯ ДЛЯ СРЕДНИХ

Пусть проверяемая гипотеза состоит в том, что параметр генеральной совокупности $\mu = a$. Например, предположим, что на основании исследований установлено предельное напряжение бесперебойного питания электронной аппаратуры. Пусть также в нашем распоряжении оказались некоторые значения напряжения, связанные с перепадами нагрузки. Делаем предположение, что перепады нагрузки не оказывают влияния на работоспособность аппаратуры, т.е. среднее значение напряжения, полученное при перепадах нагрузки, равно предельному напряжению.

Если это известное значение равно a , то нуль-гипотеза H_0 и альтернативная гипотеза H_1 записываются следующим образом: $H_0: \mu = a$, $H_1: \mu \neq a$. Примем, что переменная x в рассматриваемой совокупности имеет нормальное распределение, а величина дисперсии генеральной совокупности σ_x^2 неизвестна.

Будем проверять гипотезу H_0 с помощью статистики

$$t = \frac{\bar{x} - a}{s_x / \sqrt{n}}, \quad (7.1)$$

где s_x – выборочное стандартное отклонение.

Если $H_0: \mu = a$ справедлива, то t в выражении (7.1) имеет t -распределение Стьюдента с $n - 1$ степенями свободы. Если выбрать уровень значимости (вероятность отбросить правильную гипотезу) равным α , то в соответствии с тем, о чем шла речь в предыдущей главе, можно определить критические значения для проверки $H_0 = 0$.

В данном случае, так как распределение Стьюдента симметрично, то $(1 - \alpha)$ часть площади под кривой этого распределения с $n - 1$ степенями свободы будет заключена между точками $\pm t_{n-1} \Big|_{1 - \left(\frac{\alpha}{2}\right)}$, которые равны друг другу по абсолютной величине. Следовательно, все значения меньше отрицательного и больше положительного значения для t -распределения с заданным числом степеней свободы при выбранном уровне значимости будут составлять критическую область. Попадание выборочного значения t в эту область приводит к принятию альтернативной гипотезы.

Доверительный интервал для μ определяется из следующего выражения:

$$\bar{x} \pm t_{n-1} \frac{s_x}{\sqrt{n}} \Big|_{1 - \left(\frac{\alpha}{2}\right)}. \quad (7.2)$$

Итак, пусть известно, что предельное напряжение бесперебойного питания электронной аппаратуры равно 200,2 В. В нашем распоряжении оказалась выборка из 50 вновь найденных значений напряжений, связанных с перепадами нагрузки, для которых $\bar{x} = 220,9$ В, а $s_x = 2,25$ В. Проверим $H_0: \mu = 220,2$ против $H_1: \mu \neq 220,9$. Имеем

$$t = \frac{220,9 - 220,2}{2,25/\sqrt{50}} = 2,20.$$

Если уровень значимости выбрать $\alpha = 0,05$, то табличное критическое значение $t = \pm 2,01$ [10]. Отсюда следует, что H_0 можно отклонить в пользу H_1 на уровне значимости $\alpha = 0,05$. Таким образом, для нашего примера можно утверждать (естественно, с некоторой вероятностью), что перепады нагрузки оказывают влияния на работоспособность аппаратуры.

Замечание. t -распределение симметрично, поэтому в [10] приводятся только положительные значения критического значения t этого распределения при выбранных уровнях значимости и числе степеней свободы. Причем учитывается не только доля площади под кривой распределения справа от значения t , но и одновременно слева от значения $-t$. Это связано с тем, что в большинстве случаев при проверке гипотез нас интересует существенность отклонений сама по себе, независимо от того, в большую или меньшую сторону эти отклонения, т.е. мы проверяем $H_0: \mu = a$ против $H_1: \mu \neq a$, а не против $H_1: \mu > a$ или $H_1: \mu < a$.

Для нашего примера доверительный $100(1 - \alpha)$ – процентный интервал μ равен

$$220,9 \pm 2,01 \frac{2,25}{\sqrt{50}} = 220,9 \pm 0,64 = (220,26; 221,54).$$

Рассмотрим теперь случай, когда необходимо сравнить между собой средние двух генеральных совокупностей. Проверяемая гипотеза выглядит так: $H_0: \mu_1 - \mu_2 = 0$, $H_1: \mu_1 - \mu_2 \neq 0$. Предполагается также, что x_1 имеет нормальное распределение со средним μ_1 и дисперсией σ_x^2 , а x_2 – нормальное распределение со средним μ_2 и той же дисперсией σ_x^2 . Кроме того, принимаем, что выборки, по которым оцениваются генеральные совокупности, извлекаются независимо друг от друга и имеют объем соответственно n_1 и n_2 . Из независимости выборок следует, что если взять большее их число и для каждой пары рассчитать средние значения, то множество этих пар средних будет полностью некоррелировано.

Проверка нулевой гипотезы проводится с использованием статистики:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}, \quad (7.3)$$

где s_1^2 и s_2^2 – оценки дисперсии для первой и второй выборок соответственно. Нетрудно видеть, что (7.3) представляет собой обобщение (7.1).

Статистика (7.3) имеет t -распределение Стьюдента с $n_1 + n_2 - 2$ степенями свободы. При равенстве n_1 и n_2 , т. е. $n_1 = n_2 = n$, формула (7.3) упрощается и имеет вид

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{n}}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_{x_1}^2 + s_{x_2}^2}{n}}}. \quad (7.4)$$

ПРИМЕР 7.1. Пусть при измерении времени обработки сигнала в узле коммутации в течение двух серий испытаний получены следующие результаты: $\bar{x}_1 = 53,5$ с; $s_1^2 = 20,4$ с²; $n_1 = 485$; $\bar{x}_2 = 50,2$ с; $s_2^2 = 17,1$ с²; $n_2 = 325$. Будем считать, что условия для использования критерия Стьюдента, т.е. нормальность генеральных совокупностей, из которых взяты выборки, существование неизвестной, но одной и той же дисперсии для этих совокупностей и независимость выборок выполнены. Оценим H_0 : $\mu_1 = \mu_2$ на уровне значимости $\alpha = 0,01$. Имеем

$$t = \frac{53,5 - 50,2}{\sqrt{\frac{484 \cdot 20,4 + 324 \cdot 17,1}{485 + 325 - 2} \left(\frac{1}{485} + \frac{1}{325} \right)}} = 10,54.$$

Табличное значение $t = 2,58$ [10]. Поэтому гипотеза H_0 о равенстве средних значений времени обработки сигнала в узле коммутации в течение двух серий испытаний должна быть отвергнута на выбранном уровне значимости.

Замечание. В качестве нулевой гипотезы в математической статистике выбирается гипотеза об отсутствии значимых различий между сравниваемыми показателями, причем независимо от того, идет ли речь о средних, дисперсиях или других статистиках. И во всех этих случаях, если эмпирическое (вычисленное по формуле) значение критерия больше теоретического (выбранного из таблиц), H_0 отвергается. Если же эмпирическое значение меньше табличного, то H_0 принимается.

Для того чтобы построить доверительный интервал для разности средних $\mu_1 - \mu_2$ этих двух генеральных совокупностей, обратим внимание на то, что критерий Стьюдента, как видно из формулы (7.3), оценивает значимость разности между средними относительно стандартной ошибки этой разности. Знаменатель в (7.3) представляет именно эту стандартную ошибку. В общем случае

$$\begin{aligned} s_{x-y}^2 &= \frac{\sum_{i=1}^n ((x_i - y_i) - (\bar{x} - \bar{y}))^2}{n-1} = \frac{\sum_{i=1}^n ((x_i - \bar{x}) - (y_i - \bar{y}))^2}{n-1} = \\ &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} + \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1} - \frac{2 \sum_{i=1}^n ((x_i - \bar{x}) - (y_i - \bar{y}))}{n-1} = s_x^2 + s_y^2 - 2s_{xy}. \end{aligned}$$

Если x и y независимы, то $s_{xy} = 0$ и $s_{x-y}^2 = s_x^2 + s_y^2$.

Взяв вместо x и y выборочные значения \bar{x}_1 и \bar{x}_2 и учитывая сделанное предположение о том, что обе генеральные совокупности имеют одну и ту же дисперсию σ_x^2 , получим

$$s_{\bar{x}_1 - \bar{x}_2}^2 = \frac{\sigma_x^2}{n_1} + \frac{\sigma_x^2}{n_2} = \sigma_x^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right). \quad (7.5)$$

Оценка дисперсии σ_x^2 может быть получена из следующего соотношения:

$$\sigma_x^2 = \frac{\sum_{i=1}^{n_1} (x_{1i} - \bar{x}_1)^2 + \sum_{i=1}^{n_2} (x_{2i} - \bar{x}_2)^2}{n_1 + n_2 - 2} = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}. \quad (7.6)$$

(Знаменатель равен $n_1 + n_2 - 2$, потому что по выборкам оцениваются две величины \bar{x}_1 и \bar{x}_2 , и значит, число степеней свободы должно быть уменьшено на два.)

Если теперь подставить (7.6) в (7.5) и извлечь квадратный корень, то получится знаменатель в выражении (7.3).

Построим доверительный интервал $\mu_1 - \mu_2$ через $\bar{x}_1 - \bar{x}_2$.

Для примера получаем

$$(\bar{x}_1 - \bar{x}_2) \pm t_{n_1+n_2-2} \Big|_{1-\left(\frac{\alpha}{2}\right)} \cdot s_{\bar{x}_1 - \bar{x}_2} = 3,3 \pm 2,58 \cdot 0,313 = (2,49; 4,11).$$

Замечание. Рассмотрим предположения, используемые при построении t -критерия. Нарушения допущения о нормальности для $H_0: \mu_1 - \mu_2 = 0$ – имеют незначительное влияние на уровень значимости и мощность критерия для $n \geq 30$. Несущественно также и нарушение предположения об однородности дисперсий обеих генеральных совокупностей, из которых берутся выборки, но только в том случае, когда объемы выборок равны. Если же $n_1 \neq n_2$, а дисперсии обеих совокупностей отличаются друг от друга, то вероятности ошибок первого и второго рода будут существенно отличаться от ожидаемых.

В этом случае для проверки $H_0: \mu_1 - \mu_2 = 0$ следует пользоваться критерием

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (7.7)$$

с числом степеней свободы

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}. \quad (7.8)$$

Как правило, v получается дробным числом, поэтому при пользовании таблицами t -распределения необходимо брать табличные значения для ближайших целых значений и проводить интерполяцию для нахождения t , соответствующего полученному v .

ПРИМЕР 7.2. При изучении двух каналов связи (кабельного и спутникового) определялось время прохождения одного и того же сигнала. Были взяты две выборки с объемами $n_1 = 49$ и $n_2 = 27$. Средние и дисперсии

времени оказались равными соответственно $\bar{x}_1 = 2,34$ мс; $\bar{x}_2 = 2,08$ мс; $s_1^2 = 0,21$ мс²; $s_2^2 = 0,35$ мс². Если теперь проверять гипотезу $H_0: \mu_1 - \mu_2 = 0$ с использованием формулы (7.2), то получим, что

$$t = \frac{2,34 - 2,08}{\sqrt{\frac{48 \cdot 0,21 + 26 \cdot 0,35}{49 + 27 - 2} \left(\frac{1}{49} + \frac{1}{27} \right)}} = 2,13.$$

При уровне значимости $\alpha = 0,05$ мы должны отвергнуть нулевую гипотезу (табличное значение $t = 1,995$ для степени свободы $\nu = n_1 + n_2 - 2 = 74$ [10]) и считать, что есть статистически достоверные на выбранном уровне значимости различия между средними значениями измеряемых показателей для двух различных каналов связи.

При использовании же формул (7.7) и (7.8) имеем

$$t = \frac{2,34 - 2,08}{\sqrt{\frac{0,21}{49} + \frac{0,35}{27}}} = 1,98; \quad \nu = \frac{\left(\frac{0,21}{49} + \frac{0,35}{27} \right)^2}{\frac{(0,21/49)^2}{48} + \frac{(0,35/27)^2}{26}} = 43,5.$$

В данном случае для того же уровня значимости $\alpha = 0,05$ табличное значение $t = 2,016$ [10] и нулевая гипотеза принимается.

Этот пример показывает, что пренебрежение условиями, принимаемыми при выводе того или иного критерия, может привести к результатам прямо противоположным. Конечно же, в данном случае, имея выборки разного объема в отсутствие заранее установленного факта о том, что дисперсии измеряемого указателя в обоих каналах статистически равны, следовало пользоваться формулами (7.7) и (7.8), которые и показали отсутствие статистически значимых различий.

Поэтому проверка соблюдения всех предположений, сделанных при выводе того или иного критерия, является совершенно необходимым условием для его корректного использования.

Неизменным требованием в обеих приведенных модификациях t -критерия было требование о независимости между собой выборок. Однако на практике достаточно часто встречаются ситуации, когда это требование не может быть выполнено по объективным причинам. Например, измеряются некоторые показатели на одном и том же животном или участке территории до и после действия внешнего фактора и т.д. И в этих случаях интересует проверка гипотезы $H_0: \mu_1 - \mu_2 = 0$ против $H_1: \mu_1 - \mu_2 \neq 0$. Будем по-прежнему предполагать, что обе выборки взяты из нормальных генеральных совокупностей с одинаковой дисперсией.

В этом случае можно воспользоваться тем фактом, что разности между нормально распределенными величинами также имеют нормальное распределение, и поэтому можно воспользоваться критерием Стьюдента в форме (7.1). Таким образом, будет проверяться гипотеза о том, что n разностей

$(x_{1i} - x_{2i})$ есть выборка из нормально распределенной генеральной совокупности со средним $(\mu_1 - \mu_2)$, равным нулю.

Обозначив i -ю разность через d_i , имеем

$$t = \frac{\bar{x} - a}{s_x / \sqrt{n}} = \frac{\bar{d} - 0}{s_d / \sqrt{n}}, \quad (7.9)$$

$$\bar{d} = \frac{\sum_{i=1}^n (x_{1i} - x_{2i})}{n} = \frac{\sum_{i=1}^n d_i}{n},$$

$$s_d = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n-1}}.$$

ПРИМЕР 7.3. Пусть имеются данные о количестве импульсов, проходящих по каналу связи, за определенный интервал времени до (x_1) и после (x_2) действия помехи (табл. 7.1).

Таблица 7.1

x_1	13	9	9	9	12	8	10	11	7	9
x_2	14	13	11	14	11	11	10	13	12	12
d	-1	-4	-2	-5	1	-3	0	-2	-5	-3

Отсюда $\bar{d} = -2,4$; $s_d = 2,01$; $|t| = 3,78$. Для (7.9), имеющей t -распределение, с уровнем значимости $\alpha = 0,01$ критическое значение t для $n - 1 = 10 - 1 = 9$ степеней равно 3,25. Сравнение теоретического и эмпирического значений t -статистики показывает, что нулевая гипотеза об отсутствии статистически значимых различий между частотой импульсов до и после подачи помехи должна быть отвергнута. Можно сделать вывод о том, что помеха статистически значимо меняет частоту импульсов.

Замечание. В экспериментальных исследованиях зависимые выборки появляются достаточно часто. Этот факт иногда игнорируется, и t -критерий некорректно используется в форме (7.3).

В неправомерности этого можно убедиться, рассматривая стандартные ошибки разности между некоррелированными и коррелированными средними. В первом случае

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_x^2}{n}} = \sqrt{\sigma_x^2 \frac{2}{n}},$$

а во втором

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_x^2}{n} - 2p_{12} \frac{\sigma_x^2}{n}} = \sqrt{\sigma_x^2 \left(\frac{2}{n} - \frac{2p_{12}}{n} \right)}.$$

Стандартная ошибка разности d равна

$$s_d = s_{x_1 - x_2} = \sqrt{s_{x_1}^2 + s_{x_2}^2 - 2r_{x_1 x_2} s_{x_1} s_{x_2}}.$$

С учетом этого знаменатель в (7.9) будет иметь вид

$$\frac{s_d}{\sqrt{n}} = s_{x_1 - x_2} = \sqrt{\frac{s_{x_1}^2}{n} + \frac{s_{x_2}^2}{n} - \frac{2r_{x_1 x_2} s_{x_1} s_{x_2}}{n}}.$$

Теперь обратим внимание на то, что числители выражений (7.4) и (7.9) совпадают:

$$\bar{d} = \frac{\sum_{i=1}^n (x_{1i} - x_{2i})}{n} = \frac{\sum_{i=1}^n x_{1i}}{n} - \frac{\sum_{i=1}^n x_{2i}}{n} = \bar{x}_1 - \bar{x}_2,$$

следовательно, различие в величине t в них зависит от знаменателей.

Таким образом, если в задаче с зависимыми выборками будет использована формула (7.3) и при этом выборки будут иметь положительную корреляцию, то получаемые значения t будут меньше, чем они должны были бы быть при использовании формулы (7.9), и может возникнуть ситуация, что будет принята нулевая гипотеза, в то время как она неверна. Обратная ситуация может возникнуть, когда между выборками будет существовать отрицательная корреляция, т.е. в этом случае значимыми будут признаваться такие различия, которые на самом деле таковыми не являются.

Вернемся вновь к примеру и вычислим для приведенных данных значение t по формуле (7.3), не обращая внимания на то, что выборки связаны. Имеем: $\bar{x}_1 = 9,7$; $\bar{x}_2 = 12,1$; $s_1^2 = 3,34$; $s_2^2 = 1,88$; $|t| = 3,32$. Для числа степеней свободы, равного 18, и уровня значимости $\alpha = 0,01$ табличное значение $t = 2,88$ [10], и на первый взгляд кажется, что ничего не произошло, даже при использовании непригодной для данных условий формулы. И в этом случае вычисленное значение t приводит к отбрасыванию нулевой гипотезы, т.е. к тому же самому выводу, который был сделан с использованием правильной в данной ситуации формулы (7.9).

Однако если переформировать имеющиеся данные и представить их в виде табл. 7.2, то окажется, что это те же самые значения и они вполне могли бы быть получены в каком-нибудь из опытов. Так как все значения в обеих выборках сохранены, то использование критерия Стьюдента в формуле (7.3) дает уже полученное ранее значение $|t| = 3,32$ и приводит к тому же самому выводу, который уже был сделан.

Таблица 7.2

x_1	13	9	12	9	10	9	11	7	9	8
x_2	10	13	12	12	11	13	11	14	11	14

А теперь рассчитаем значение t по формуле (7.9), которая и должна использоваться в данном случае. Имеем: $\bar{d} = -2,4$; $s_d = 3,03$; $|t| = 2,51$. Критическое значение t при выбранном уровне значимости и девяти степенях свободы равно 3,25. Следовательно, оснований отвергнуть нулевую гипотезу нет, она принимается. Оказывается, что этот вывод прямо противоположен тому, который был сделан при использовании формулы (7.3).

На этом примере видно, как важно для получения правильных выводов при анализе экспериментальных данных строго соблюдать все требования, которые были положены в основу определения того или иного критерия.

Рассмотренные модификации критерия Стьюдента предназначены для проверки гипотез относительно средних двух выборок. Однако возникают ситуации, когда появляется необходимость сделать выводы относительно равенства одновременно k средних. Для этого случая применяется дисперсионный анализ.

7.2. КРИТЕРИИ СОГЛАСИЯ ДЛЯ ДИСПЕРСИЙ

Проверка статистических гипотез относительно дисперсий генеральных совокупностей проводится в той же последовательности, что и для средних.

1. Формулируется нулевая гипотеза (об отсутствии статистически значимых различий между сравниваемыми дисперсиями).

2. Делаются некоторые предположения относительно выборочного распределения статистики, с помощью которой планируется оценивать параметр, входящий в гипотезу.

3. Выбирается уровень значимости для проверки гипотезы.

4. Вычисляется значение интересующей статистики и принимается решение относительно истинности нулевой гипотезы.

Проверяется гипотеза о том, что дисперсия генеральной совокупности $\sigma_x^2 = a$, т.е. $H_0: \sigma_x^2 = a$ против $H_1: \sigma_x^2 \neq a$. Если предположить, что переменная x имеет нормальное распределение и что выборка объемом n извлекается из генеральной совокупности случайно, то для проверки нулевой гипотезы H_0 используется статистика

$$\chi^2 = \frac{(n-1)s_x^2}{a}. \quad (7.10)$$

Подставляя формулу для дисперсии в (7.10), получаем

$$\chi^2 = \frac{(n-1) \sum_{i=1}^n (x_i - \bar{x})^2}{a(n-1)} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{a}. \quad (7.11)$$

Из этого выражения видно, что числитель представляет собой сумму квадратов отклонений нормально распределенных величин от их среднего. Каждое из этих отклонений также распределено нормально. Поэтому в соответствии с распределением суммы квадратов нормально распределенных величин статистики (7.10) и (7.11) имеют χ^2 -распределение с $n - 1$ степенью свободы.

По аналогии с использованием t -распределения при проверке гипотезы $H_0: \sigma_x^2 = a$ для выбранного уровня значимости α по таблице распределения χ^2 [10] устанавливаются критические точки, соответствующие вероятностям принятия нулевой гипотезы $1 - \alpha/2$ и $\alpha/2$. Доверительный интервал для σ_x^2 при выбранном α строится следующим образом:

$$\frac{(n-1)s_x^2}{\chi_{n-1}^2 \Big|_{1-\frac{\alpha}{2}}} < \sigma_x^2 < \frac{(n-1)s_x^2}{\chi_{n-1}^2 \Big|_{\frac{\alpha}{2}}}. \quad (7.12)$$

ПРИМЕР 7.4. Пусть на основании обширных экспериментальных исследований установлено, что дисперсия надежности электронного оборудования определенной серии равна 4,37 условных единиц. В распоряжение специалиста попадает выборка объемом $n = 28$ такого оборудования, предположительно той же серии. Проведенный анализ показал, что для этой выборки $s_x^2 = 5,01$, и нужно убедиться в том, что эта выборка и известная ранее дисперсия статистически неразличимы на уровне значимости $\alpha = 0,1$.

По формуле (7.10) имеем

$$\chi^2 = \frac{(28-1) \cdot 5,01}{4,37} = 30,95.$$

Полученную величину необходимо сравнить с критическими значениями $\alpha/2 = 0,05$; и $(1 - \alpha/2) = 0,95$. Из таблицы для χ^2 -распределения [10] с 27 степенями свободы имеем соответственно 40,1 и 16,2, откуда следует, что нулевая гипотеза может быть принята. Соответствующий доверительный интервал по формуле (7.12) для s_x^2 равен $3,37 < \sigma_x^2 < 8,35$.

В отличие от проверки гипотез относительно выборочных средних с использованием критерия Стьюдента, когда ошибки первого и второго рода несущественно менялись при нарушении предположения о нормальном распределении генеральных совокупностей, в случае гипотез о дисперсиях при невыполнении условий нормальности ошибки меняются существенно.

Рассмотренная выше задача о равенстве дисперсии некоторому фиксированному значению представляет ограниченный интерес, так как довольно редко встречаются ситуации, когда известна дисперсия генеральной совокупности. Значительно больший интерес представляет случай, когда нужно проверить, равны ли дисперсии двух совокупностей, т.е. проверка гипотезы $H_0: \sigma_1^2 = \sigma_2^2$ против альтернативы $H_1: \sigma_1^2 \neq \sigma_2^2$. При этом предполагается, что выборки объемом n_1 и n_2 случайно извлекаются из генеральных совокупностей с дисперсиями σ_1^2 и σ_2^2 .

Для проверки нулевой гипотезы используется критерий отношения дисперсий Фишера

$$F = \frac{s_1^2}{s_2^2}. \quad (7.13)$$

Поскольку суммы квадратов отклонений нормально распределенных случайных величин от их средних значений имеют распределение χ^2 , то и числитель и знаменатель (7.13) представляют собой величины с распределением χ^2 , поделенные соответственно на n_1 и n_2 , и, следовательно, их отношение имеет F -распределение с $n_1 - 1$ и $n_2 - 1$ степенями свободы.

Общепринято (и так построены таблицы F -распределения), что в качестве числителя в (7.13) берется большая из дисперсий, и поэтому определяется только одна критическая точка, соответствующая выбранному уровню значимости.

Пусть в нашем распоряжении оказались две выборки объемом $n_1 = 11$ и $n_2 = 28$ из различных серий оборудования, для которых надежности имеют дисперсии $s_1^2 = 0,59$ и $s_2^2 = 0,38$. Необходимо проверить гипотезу о равенстве дисперсий для экспериментальных серий оборудования при уровне значимости $\alpha = 0,05$. Имеем

$$F_{10,27} = \frac{0,59}{0,38} = 1,552.$$

Из таблицы F -распределения [10] критическое значение с (10,27) степенями свободы равно 2,20. Следовательно, нулевая гипотеза принимается.

Замечание. В литературе иногда можно встретить утверждение о том, что проверке гипотезы о равенстве средних по критерию Стьюдента должна предшествовать проверка гипотезы о равенстве дисперсий. Это неправильная рекомендация. Более того, она может привести к ошибкам, которых можно избежать, если ей не следовать.

В самом деле, результаты проверки гипотезы о равенстве дисперсий с использованием критерия Фишера в значительной мере зависят от предположения о том, что выборки взяты из совокупностей с нормальным распределением. В то же время критерий Стьюдента малочувствителен к нарушениям нормальности, и если удастся получить выборки равного объема, то предположение о равенстве дисперсий также не является существенным. В случае неравных n следует пользоваться для проверки гипотезы формулами (7.7) и (7.8).

При проверке гипотез о равенстве дисперсий возникают некоторые особенности в расчетах, связанных с зависимыми выборками. В этом случае для проверки гипотезы $H_0: \sigma_1^2 = \sigma_2^2$ против альтернативы $H_1: \sigma_1^2 \neq \sigma_2^2$ используется статистика

$$t = \frac{s_1^2 - s_2^2}{\sqrt{\frac{4s_1^2 s_2^2}{n-2}(1-r_{12}^2)}}. \quad (7.14)$$

Если нулевая гипотеза справедлива, то статистика (7.14) имеет t -распределение Стьюдента с $n - 2$ степенями свободы.

При измерении 35 скачков напряжения была получена дисперсия $s_1^2 = 134,5$. Повторные измерения через две недели показали $s_2^2 = 199,1$. При этом коэффициент корреляции между парными измерениями оказался равным $r_{12} = 0,876$. Если не обращать внимание на то, что выборки зависимы, и воспользоваться критерием Фишера для проверки гипотезы $H_0: \sigma_1^2 = \sigma_2^2$, то получим $F = 1,48$. Если выбрать уровень значимости $\alpha = 0,05$, то нулевая

гипотеза будет принята, так как критическое значение F -распределения для $f_1 = 35 - 1 = 34$ и $f_2 = 35 - 1 = 34$ степеней свободы равно 1,79 [10].

В то же время, если использовать подходящую для данного случая формулу (7.14), то получим $t = 2,35$, в то время как критическое значение t для 33 степеней свободы и выбранного уровня значимости $\alpha = 0,05$ равно $\approx 2,03$ [10]. Следовательно, нулевая гипотеза о равенстве дисперсий в этих двух выборках должна быть отклонена. Таким образом, из этого примера видно, что, как и в случае проверки гипотезы о равенстве средних, использование критерия, не учитывающего специфику экспериментальных данных, приводит к ошибке.

Существует критерий Бартлетта, используемый при проверке гипотез об одновременном равенстве k дисперсий. Кроме того, что вычисления статистики этого критерия довольно трудоемки, основной недостаток этого критерия в том, что он необычайно чувствителен к отклонениям от предположения о нормальности распределений совокупностей, из которых извлекаются выборки. Таким образом, при его использовании никогда нельзя быть уверенным в том, что нулевая гипотеза отклонена из-за того, что статистически значимо различаются дисперсии, а не из-за того, что выборки не имеют нормального распределения. Поэтому в случае возникновения проблемы сравнения нескольких дисперсий необходимо искать такую постановку задачи, когда можно будет использовать критерий Фишера или его модификации.

7.3. КРИТЕРИИ СОГЛАСИЯ ДЛЯ КОРРЕЛЯЦИОННЫХ ПОКАЗАТЕЛЕЙ

Проверяется гипотеза о том, что обычный коэффициент ρ_{xy} корреляции равен некоторому фиксированному значению a , т.е. $H_0: \rho_{xy} = a$ против альтернативной гипотезы $H_1: \rho_{xy} \neq a$.

При этом предполагается, что случайная выборка n пар наблюдений извлекается из двумерной нормальной генеральной совокупности, в которой корреляция между x и y численно равна ρ_{xy} . Для проверки H_0 используется статистика

$$z = \frac{z_r - z_a}{1/\sqrt{n-3}}, \quad (7.15)$$

где z_r и z_a – z -преобразованные по Фишеру (6.23) значения, соответствующие выборочному r_{xy} и a .

Как правило, проверяется гипотеза о том, что $\rho_{xy} = 0$. В этом случае статистика (7.15) приобретает вид

$$z = \frac{z_r}{1/\sqrt{n-3}}. \quad (7.16)$$

Если нулевая гипотеза верна, то статистика (7.16) имеет единичное нормальное распределение и способ построения доверительных интервалов для z_r , а следовательно и для ρ_{xy} , очевиден.

ПРИМЕР 7.5. Пусть для данных о ежедневном доходе на маршруте трамвая №13 и среднедневной температуре воздуха в зимний месяц за 28 дней был рассчитан коэффициент корреляции, который оказался равным $r_{xy} = 0,38$.

Проверим с уровнем значимости $\alpha = 0,05$ гипотезу $H_0: \rho_{xy} = a$. По формуле (6.23) получим величину z -преобразования $z_r = 0,4$. Отсюда $z = 0,4\sqrt{28-3} = 2,0003$. Из таблицы нормального распределения [10] следует, что критической точкой для $\alpha = 0,05$ является величина $\pm 1,96$. Следовательно, на данном уровне значимости нулевая гипотеза должна быть отвергнута.

Для построения доверительного интервала сначала вычисляют

$$z_r \pm z \frac{1}{\sqrt{n-3}} \Big|_{1-\left(\frac{\alpha}{2}\right)} = 0,40 \pm (1,96/\sqrt{28-3}) = 0,4 \pm 0,392.$$

Полученные две точки на шкале z -преобразования (0,008; 0,792) переводят в значения r_{xy} по формуле

$$r = \frac{e^{2z} - 1}{e^{2z} + 1} \quad (7.17)$$

и получают доверительный интервал для ρ_{xy} с выбранным уровнем значимости $0,008 < \rho_{xy} < 0,659$.

Кроме статистики (7.16) для проверки существует еще одна статистика $H_0: \rho_{xy} = a$

$$t = \frac{r_{xy}}{\sqrt{(1-r_{xy}^2)/(n-2)}}, \quad (7.18)$$

которая при $\rho_{xy} = 0$ имеет t -распределение с $n - 2$ степенями свободы. Использование этой статистики позволило составить таблицы критических значений r_{xy} , превышение которых приводит к отклонению $H_0: \rho_{xy} = 0$. Для примера при $r_{xy} = 0,38$ расчетное $t = 2,095$, а табличное значение для $\alpha = 0,05$ равно 2,069. Следовательно, на данном уровне значимости нулевая гипотеза должна быть отвергнута.

Проверка статистических гипотез о равенстве двух коэффициентов корреляции, т.е. $H_0: \rho_1 = \rho_2$ против $H_1: \rho_1 \neq \rho_2$, также проводится с использованием статистики

$$z = \frac{z_{r_1} - z_{r_2}}{\sqrt{\frac{1}{\sqrt{n_1-3}} + \frac{1}{\sqrt{n_2-3}}}}, \quad (7.19)$$

имеющей единичное нормальное распределение. Это справедливо, если выборки извлекаются из двумерных нормальных генеральных совокупностей, характеризующихся коэффициентами корреляции ρ_1 и ρ_2 .

При вычислениях по формуле (7.19) последовательность действий аналогична приведенной выше.

ПРИМЕР 7.6. При подсчете коэффициента корреляции с $n_1 = 14$ и $n_2 = 19$ были получены значения $r_1 = -0,785$ и $r_2 = -0,624$. Необходимо проверить гипотезу о равенстве коэффициентов корреляции в совокупностях, из которых взяты эти выборки, на уровне значимости $\alpha = 0,01$.

Имеем $z_{r1} = -1,06$, $z_{r2} = -0,73$. (Знак z -преобразованной величины зависит от знака коэффициента корреляции, а абсолютная величина остается одной и той же для равных по величине, но имеющих разные знаки коэффициентов корреляции.)

По формуле (7.19) подсчитываем

$$z = \frac{-1,06 - (-0,73)}{\sqrt{\frac{1}{14-3} + \frac{1}{19-3}}} = -0,834.$$

Отсюда следует, что нулевая гипотеза принимается, так как критическая точка для $\alpha = 0,01$ в единичном нормальном распределении равна $\pm 2,58$.

Замечание. Нулевая гипотеза, которая проверялась, утверждает, что $\rho_{r1} = \rho_{r2}$. Эта гипотеза имеет определенный смысл и всегда может быть проверена. В то же время, имея в виду формулу (7.19), иногда утверждают, что с ее помощью проверяют гипотезу о разности между r_1 и r_2 . Это некорректное утверждение, потому что рассмотрение этой разности самой по себе лишено смысла. В самом деле, разность, равную, например, 0,3, можно получить для пар коэффициентов корреляции, равных 0,9 и 0,6, 0,5 и 0,2, 0,1 и -0,2 и т.д. Естественно, что в каждом из этих случаев значимость разности 0,3 существенно различна, потому что она отражает неодинаковые возможности предсказания одной переменной по другой. Это следует помнить, чтобы правильно формулировать проверяемую гипотезу, а затем проводить интерпретацию полученного вывода в содержательных понятиях.

Гипотезы о равенстве двух ранговых или двух частных коэффициентов корреляции могут быть также проверены с использованием статистики (7.18) или соответствующей ей таблицы с той только разницей, что для ранговых коэффициентов корреляции вместо r_{xy} используются значения r_s , а при проверке частных коэффициентов корреляции вместо r_{xy} подставляются значения вида $r_{xy.z}$, а в знаменателе вместо делителя $n - 2$ используется делитель $n - 3$.

7.4. КРИТЕРИИ ДЛЯ ПРОВЕРКИ СЛУЧАЙНОСТИ И ОЦЕНКИ РЕЗКО ВЫДЕЛЯЮЩИХСЯ НАБЛЮДЕНИЙ

Довольно часто данные получают сериями во времени или пространстве. Например, при определении фактических показателей точности технологического процесса, установлении соответствия качества продукции требованиям нормативно-технической документации, сравнении различных образцов продукции и т.д. Являются ли полученные данные независимыми? Такое предположение обеспечивает достоверность при вычислении различных статистик. Поэтому требуется его проверка.

Рассмотрим критерий для проверки нулевой гипотезы о независимости одинаково нормально распределенных величин. Этот критерий является параметрическим. Он основан на расчете среднего квадратов последовательных разностей

$$d^2 = \frac{1}{n-1} \sum_{i=1}^{n-1} (x_{i+1} - x_i)^2. \quad (7.20)$$

Если ввести новую статистику $q = d^2/2s^2$, то при справедливости нулевой гипотезы статистика

$$A = (q-1) \sqrt{\frac{n^2-1}{n-2}} \quad (7.21)$$

для $n > 10$ распределена асимптотически по стандартному нормальному распределению.

ПРИМЕР 7.7. В табл. 7.3 приведены времена реакции (x_i) испытуемого в одном из психофизиологических экспериментов. Величина d_i в таблице соответствует разности между двумя соседними значениями x_i .

Таблица 7.3

i	x_i	d_i	i	x_i	d_i
1	630		13	690	15
2	610	-20	14	605	-85
3	680	70	15	580	-25
4	695	15	16	595	15
5	645	-50	17	650	55
6	675	35	18	625	-25
7	605	-70	19	630	5
8	595	-10	20	660	30
9	705	110	21	595	-65
10	680	-25	22	615	20
11	640	-40	23	645	30
12	675	35			

Имеем: $s^2 = 1364,7$; $d^2 = 2176,1$; $q = 0,797$, откуда

$$A = (0,797 - 1) \sqrt{\frac{23^2 - 1}{23 - 2}} = -1,02.$$

Поскольку для нормального распределения при $\alpha = 0,05$ критическое значение равно $\pm 1,96$, нулевая гипотеза о независимости полученного ряда принимается с выбранным уровнем значимости.

Другой вопрос, который часто возникает при анализе экспериментальных данных: что делать с некоторыми наблюдениями, которые резко отличаются от основной массы наблюдений. Такие резко выделяющиеся наблюдения могут возникнуть при методических ошибках, ошибках вычислений и т.д. Во всех тех случаях, когда экспериментатору известно, что в наблюдение вкралась ошибка, он должен исключать это значение независимо от его величины. В других случаях существует только подозрение на ошибку, и тогда необходимо использовать соответствующие критерии, с тем, чтобы принять то или иное решение, т.е. исключить или оставить резко выделяющиеся наблюдения.

В общем случае вопрос ставится так: наблюдения произведены над одной и той же генеральной совокупностью или некоторая часть или отдельные значения относятся к другой генеральной совокупности?

Конечно, единственно надежным способом для исключения отдельных наблюдений является тщательное изучение условий, при которых эти наблюдения получены. Если по каким-то причинам условия отличались от стандартных, то наблюдения должны быть исключены из дальнейшего анализа с использованием критериев.

Приведем статистики, которые могут быть использованы для проверки гипотезы о том, что наблюдения производятся случайно над одной и той же генеральной совокупностью

$$\lambda = \frac{x_{(n)} - x_{(n-1)}}{s_x}, \quad (7.22)$$

$$\lambda^* = \frac{x_{(n)} - \bar{x}}{s'_x}, \quad (7.23)$$

$$z_n = \sum_{i=1}^n \frac{x_i}{x_{(n)}}, \quad (7.24)$$

где $x_{(n)}$ – подозреваемое на «выброс» наблюдение. Если все значения ряда проранжировать, то в нем резко выделяющееся наблюдение будет занимать n -е последнее место.

Для статистики (7.22) протабулирована функция распределения. В табл. 7.4 приведены критические точки этого распределения для некоторых n .

Таблица 7.4

n	$\lambda_{0,05}$	$\lambda_{0,01}$
10	1,5	2,0
20	1,3	1,8
30	1,2	1,7
50	1,1	1,6
100	1,0	1,5
1000	0,8	1,2

Критическими значениями для статистики (7.23) в зависимости от n являются

$$\begin{aligned} \lambda^* &= 4,0; & 6 < n < 100, \\ \lambda^* &= 5,6; & 100 < n < 1000, \\ \lambda^* &= 5,0; & n > 1000. \end{aligned}$$

В формуле (7.23) предполагается, что \bar{x} и s'_x вычисляются без учета подозреваемого наблюдения.

Для статистики (7.24) необходимо знать математическое ожидание и дисперсию. В частном случае, если x_i распределены равномерно, то математическое ожидание и дисперсия z_n имеют вид

$$M[z_n] = \frac{n+1}{2}, \quad D[z_n] = \frac{n-1}{12}.$$

Критическую область образуют малые значения z_n , которые соответствуют большим значениям $x_{(n)}$. Если интересует проверка на «выброс» наименьшего значения $x_{(1)}$, то сначала преобразуют данные, чтобы они имели равномерное распределение на интервале $[0,1]$, а затем берут дополнение этих равномерных величин до 1 и проверяют $x_{(n)}$ по формуле (7.24).

ПРИМЕР 7.8. Рассмотрим использование приведенных критериев для следующего проранжированного ряда наблюдений: 3, 4, 5, 5, 6, 7, 8, 9, 9, 10, 11, 17. Допустим, данная выборка представляет равномерно распределенную величину. Необходимо решить, следует ли отвергнуть наибольшее значение 17. Имеем: $\bar{x} = 7,0$; $s_x = 3,81$; $s'_x = 2,61$. По формуле (7.22) $\lambda = (17 - 11)/3,81 = 1,57$, и нулевая гипотеза должна быть принята при $\alpha = 0,01$. По формуле (7.23) $\lambda^* = (17 - 7,0)/2,61 = 3,83$, и нулевая гипотеза также должна быть принята. Для использования третьего критерия найдем $z_n = 5,53$. Определим нормированную статистику

$$w = \frac{z_n - M[z_n]}{\sqrt{D[z_n]}} = \frac{z_n - \frac{n+1}{2}}{\sqrt{\frac{n-1}{12}}} = \frac{5,53 - 6,5}{\sqrt{11/12}} = -1,013.$$

Статистика распределена нормально с нулевым средним и единичной дисперсией. Критическое значение при $\alpha = 0,05$ равно $\pm 2\sigma$. Следовательно, нулевая гипотеза о том, что наибольшее значение 17 нельзя отвергнуть, принимается.

8. ОСНОВЫ ДИСПЕРСИОННОГО АНАЛИЗА

Дисперсионный анализ является одним из наиболее распространенных методов, используемых для сравнения большого числа наборов случайных величин. Он широко применяется в планировании экспериментов, для проверки гипотез о наличии связи между откликом (ответом) системы и исследуемыми факторами, а также для установления силы влияния факторов и их взаимодействий.

Основная идея метода заключается в изучении источников изменчивости зависимой переменной (отклика) и разложении общей дисперсии наблюдаемых значений отклика на составляющие – дисперсию, обусловленную влиянием изучаемых факторов, и остаточную дисперсию, являющуюся следствием действия случайных причин и неучтенных факторов. Сравнивая дисперсию факторов с остаточной дисперсией, можно проверить гипотезу о влиянии факторов на величину изменчивости отклика [10].

Можно сказать, что дисперсионный анализ представляет собой систему понятий и технических приемов, позволяющих обобщить процедуру сравнения двух средних для двух выборок, взятых из генеральных совокупностей с нормальным распределением, на случай большого числа выборок.

8.1. ДВЕ МОДЕЛИ ДИСПЕРСИОННОГО АНАЛИЗА

Рассмотрим два простых примера.

ПРИМЕР 8.1. Пусть из данной продукции – прибора высокой точности – берутся kn деталей. Затем k групп из n деталей каждая обрабатывается на отдельном станке, и после обработки проверяется качество каждой детали.

В этом примере мы рассматриваем качество в n деталях системы, извлеченной в качестве случайной выборки из нормально распределенной генеральной совокупности со средним значением и дисперсией, которые характеризуют эффект специфических условий обработки этих деталей. Если бы мы взяли другую случайную выборку в n деталей, обработали ее на тех же отдельных станках и провели те же измерения, то можно было бы считать, что мы имеем вторую выборку, извлеченную из той же самой генеральной совокупности и т. д. Иные системы наблюдений, соответствующие другим условиям обработки, интерпретируются аналогичным образом, т. е. каждому из k условий обработки сопоставляется нормально распределенная генеральная совокупность, и при этом предполагается, что k систем данных о качестве деталей извлечены из этих совокупностей случайным образом. Генеральные совокупности характеризуется k математическими ожиданиями и k дисперсиями, которые оцениваются k выборочными средними и дисперсиями соответственно.

Если эксперимент необходимо повторить, то создаются те же самые k условий обработки деталей, так как цель такого эксперимента состоит именно в анализе эффектов выбранных условий.

Таким образом, определяющим моментом этого типа дисперсионного анализа является предположение о том, что мы имеем k нормально распределенных генеральных совокупностей с математическими ожиданиями $\mu_1, \mu_2, \dots, \mu_k$ и одной и той же дисперсией σ^2 . Это предположение содержит $k + 1$ неизвестных параметров, а именно k математических ожиданий и одну дисперсию. Целью дисперсионного анализа этого типа является оценка этих параметров, отыскание выборочных распределений оценок и проверка определенных гипотез относительно значений параметров. В частности, в приведенном примере в первую очередь проверяется гипотеза о том, что k генеральных средних равны, т. е. различия в условиях обработки не влияют на качество деталей прибора.

Главной особенностью рассматриваемого типа дисперсионного анализа является то, что каждое наблюдение рассматривается как сумма двух компонент, а именно неизвестного среднего значения генеральной совокупности, из которой получено наблюдение и которое одинаково для всех наблюдений в пределах системы, и отклонения от этого среднего значения. Первая компонента считается неизвестной постоянной, тогда как вторая считается случайной величиной. Систематические различия между генеральными совокупностями характеризуются k генеральными средними, которые выражают эффекты k причин (способов обработки, условий хранения и т. д.).

Рассмотрим второй пример.

ПРИМЕР 8.2. Из продукции производства окрашенных листов полистирола случайным образом берутся k листов. Из каждого листа случайным образом вырезается n кусков, и определяется толщина окрашенного слоя.

Из каждого листа мы берем случайным образом n кусков и определяем толщину окраски. Эти измерения отклоняются случайным образом, потому что невозможно обеспечить одинаковую глубину проникновения красителя по всей поверхности листа из-за неоднородностей, которые всегда имеют место. Предполагается, что n измерений образуют случайную выборку, извлеченную из гипотетической бесконечной совокупности нормально распределенных измерений глубины проникновения красителя в одном и том же листе. Для i -го листа среднее значение соответствующей генеральной совокупности μ_i является мерой глубины проникновения красителя в этом листе, а дисперсия σ^2 является мерой неоднородности листа. (Предполагается, что дисперсия одна и та же для всех листов.)

Из-за изменений в технологическом процессе глубина проникновения красителя может изменяться от листа к листу. Отсюда следует, что k листов рассматриваются как случайная выборка, извлеченная из гипотетической бесконечной совокупности окрашенных листов, которые могли быть изготовлены в данном процессе. Предполагая, что распределение глубины проникновения красителя в листы полистирола для данной совокупности нормально, можно рассматривать систему μ_i как случайную выборку из этой совокупности, которая характеризуется генеральным средним μ и дисперсией σ^2 .

Для данного примера повторение эксперимента означает, что берется новая выборка из k окрашенных листов, потому что конечной целью является характеристика процесса окрашивания в целом, а не сравнение каких-то конкретных k листов.

Главная особенность этого вида дисперсионного анализа состоит в предположении о том, что каждое измерение глубины проникновения красителя имеет две случайные компоненты:

- отклонение измерения от μ_i ;
- отклонение μ_i от μ .

Такое предположение влечет за собой появление трех неизвестных параметров – среднего и двух дисперсий, а целью проводимого дисперсионного анализа по-прежнему является оценка этих параметров, отыскание их выборочных распределений, проверка определенных гипотез относительно значений параметров.

В данном примере интерес представляет сравнение всех трех параметров, а затем сравнение двух дисперсий с тем, чтобы по результату такого сравнения определить, какой источник рассеяния является более важным. Таким образом, основным в этом виде дисперсионного анализа является то, что каждое измерение рассматривается как сумма уже не двух, как в предыдущем случае, а трех компонент:

- неизвестной постоянной – среднего значения, одного и того же для всех измерений;
- случайной компоненты, связанной с группой причин, приводящих к рассеянию внутри систем измерений;
- случайной компоненты, связанной с группой причин, приводящих к рассеянию между системами измерений.

Конечно, существуют значительно более сложные, чем рассмотренные, варианты дисперсионного анализа, зависящие одновременно и от множества факторов и от комбинаций постоянных и случайных эффектов, однако общие идеи, связанные с четким разграничением систематических и случайных компонент и существованием показателей изменчивости внутри и между системами наблюдений или измерений, остаются одними и теми же.

8.2. РАЗБИЕНИЕ СУММЫ КВАДРАТОВ

Рассмотрим только один фактор с несколькими уровнями. Результаты экспериментов или наблюдений обычно представляются, как в табл. 8.1.

Таблица 8.1

Номер опыта	Уровень фактора				
	1		j		k
1	x_{11}		x_{1j}		x_{1k}
...					
i	x_{i1}		x_{ij}		x_{ik}
n	x_{n1}		x_{nj}		x_{nk}

По результатам применения каждого уровня действующего фактора вычисляются средние значения \bar{x}_j . Однако сравниваются между собой не полученные средние, а так называемые эффекты $\alpha_j = \bar{x}_j - \bar{x}$, где \bar{x} – среднее всех наблюдений. Тогда результат i -го опыта при использовании уровня фактора j можно записать в виде

$$x_{ij} = \bar{x}_j + e_{ij} = \bar{x} + \alpha_j + e_{ij}, \quad (8.1)$$

где e_{ij} – ошибка результата.

Из этого равенства следует, что

$$x_{ij} - \bar{x} = \alpha_j + e_{ij}. \quad (8.2)$$

Возведя обе части в квадрат, получим

$$(x_{ij} - \bar{x})^2 = \alpha_j^2 + 2\alpha_j e_{ij} + e_{ij}^2. \quad (8.3)$$

Таких равенств для каждого уровня j всего n , а всего уровней – k . Поэтому складывая все равенства такого вида для каждого уровня, получим

$$\sum_{j=1}^k \sum_{i=1}^n (x_{ij} - \bar{x})^2 = \sum_{j=1}^k \sum_{i=1}^n \alpha_j^2 + 2 \sum_{j=1}^k \sum_{i=1}^n \alpha_j e_{ij} + \sum_{j=1}^k \sum_{i=1}^n e_{ij}^2. \quad (8.4)$$

С учетом того, что $\sum_{j=1}^k \sum_{i=1}^n \alpha_j e_{ij} = \sum_{j=1}^k \left(\alpha_j \sum_{i=1}^n e_{ij} \right) = 0$,

$$\sum_{j=1}^k \sum_{i=1}^n (x_{ij} - \bar{x})^2 = \sum_{j=1}^k \sum_{i=1}^n \alpha_j^2 + \sum_{j=1}^k \sum_{i=1}^n e_{ij}^2$$

или

$$\sum_{j=1}^k \sum_{i=1}^n (x_{ij} - \bar{x})^2 = \sum_{j=1}^k \sum_{i=1}^n (\bar{x}_j - \bar{x})^2 + \sum_{j=1}^k \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2.$$

Это тождество и является основным в дисперсионном анализе. Из него следует, что варьирование всех результатов наблюдений или измерений около общего среднего может быть разложено на суммы квадратов, первая из которых характеризует варьирование, обусловленное изменчивостью эффектов различных уровней фактора, а вторая – варьирование под влиянием неучтенных факторов (ошибка эксперимента).

8.3. ОДНОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ С ПОСТОЯННЫМИ ЭФФЕКТАМИ

Рассмотрим базовый вариант дисперсионного анализа с постоянными эффектами. Нас будет интересовать процесс статистического вывода относительно множества главных эффектов $\alpha_1, \alpha_2, \dots, \alpha_k$.

Пусть имеется бесконечная последовательность повторений эксперимента, в каждом из которых будет n объектов с k уровнями. При этом k совокупностей фиксируется таким образом, чтобы каждая из них обеспечивала получение одной выборки объемом n при каждом повторении эксперимента.

Таким образом, мы приходим к схеме дисперсионного анализа с постоянными эффектами. Будем также предполагать, что n наблюдений на каждом уровне независимы и взяты из нормальной генеральной совокупности с дисперсией σ^2 . Предполагается, что дисперсия σ^2 одинакова на всех k уровнях, а число наблюдений для них одинаково и равно n .

Дисперсионный анализ включает в себя следующие этапы:

1. Для объяснения данных постулируется модель $x_{ij} = \bar{x} + \alpha_j + e_{ij}$, где все e_{ij} независимы, а $\alpha_1 + \alpha_2 + \dots + \alpha_k = 0$.

2. Формулируется нуль-гипотеза $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_k$ против H_1 , что, по крайней мере, два α_i различны. Или иначе

$$H_0: \bar{x}_1 = \bar{x}_2 = \dots = \bar{x}_k, H_1: \sum_{j=1}^k (\bar{x}_j - \bar{x})^2 \neq 0.$$

3. Выбирается уровень значимости α , как правило, между 0,01 и 0,1.

4. Производятся вычисления сумм квадратов ($СК$), степеней свободы и средних квадратов ($ССК$) по следующим формулам: между уровнями

$$СК_M = \sum_{j=1}^k n (\bar{x}_j - \bar{x})^2 = \frac{\sum_j \left(\sum_i x_{ij} \right)^2}{n} - \frac{\left(\sum_j \sum_i x_{ij} \right)^2}{kn},$$

внутри уровней

$$CK_B = \sum_{j=1}^k \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2 = \sum_{j=1}^k \sum_{i=1}^n x_{ij}^2 - \frac{\sum_j \left(\sum_i x_{ij} \right)^2}{n},$$

при этом число степеней свободы для CK_M равно $(k-1)$, а для $CK_B - k(n-1)$ и $ССК_M = \frac{CK_M}{(k-1)}$, $ССК_B = \frac{CK_B}{k(n-1)}$.

5. Рассчитывается дисперсионное отношение $F = ССК_M / ССК_B$ и сравнивается с критическим значением F -распределения для выбранного уровня значимости.

ПРИМЕР 8.3. В табл. 8.2 приведены данные о времени простой зрительно-моторной реакции 40 испытуемых на предъявляемые стимулы разной интенсивности [10]. Будем проверять гипотезу $H_0: \bar{x}_1 = \bar{x}_2 = \bar{x}_3 = \bar{x}_4$ с уровнем значимости $\alpha = 0,05$.

Таблица 8.2

Уровень интенсивности стимула			
1	2	3	4
260	510	520	410
340	500	640	490
460	330	390	560
480	280	540	640
420	470	580	720
490	500	530	650
740	480	770	630
610	600	560	870
510	710	630	770
530	420	590	620

Найдем степени свободы для $ССК_M$ и $ССК_B$, которые будут равны соответственно $4 - 1 = 3$ и $4(10 - 1) = 36$. Промежуточные данные приведены в табл. 8.3.

Таблица 8.3

Уровень интенсивности стимула			
1	2	3	4
$n_1=10$	$n_2=10$	$n_3=10$	$n_4=10$
$\bar{x}_1 = 484$	$\bar{x}_2 = 480$	$\bar{x}_3 = 575$	$\bar{x}_4 = 636$
$\sum x_{i1} = 4840$	$\sum x_{i2} = 4800$	$\sum x_{i3} = 5750$	$\sum x_{i4} = 6360$
$\sum x_{i1}^2 = 2502400$	$\sum x_{i2}^2 = 2439200$	$\sum x_{i3}^2 = 3392500$	$\sum x_{i4}^2 = 4203400$

Используя эти данные, имеем

$$\sum_j \sum_i x_{ij} = 21750, \quad \sum_j \sum_i x_{ij}^2 = 12537500;$$

$$CK_M = \frac{(4840)^2 + (4800)^2 + (5750)^2 + (6360)^2}{10} - \frac{(21750)^2}{40} = 171208;$$

$$CK_B = 12537500 - \frac{(4840)^2 + (4800)^2 + (5750)^2 + (6360)^2}{10} = 539730;$$

$$ССК_M = \frac{171208}{3} = 57069; \quad ССК_B = \frac{539730}{36} = 14993; \quad F = \frac{57069}{14993} = 3,806.$$

Критическое значение F -распределения для 3 и 36 степеней свободы при выбранном уровне значимости $\alpha = 0,01$ равно 4,378. Поэтому нулевая гипотеза на данном уровне значимости должна быть принята. Делаем вывод, что использование разных интенсивностей стимула не приводит к статистически значимым изменениям времени реакции.

Достаточно часто на практике сталкиваются с ситуацией, когда в силу тех или иных причин число наблюдений n_j для каждого уровня фактора неодинаково. Изложенная выше методика распространяется и на этот случай. Число степеней свободы для $ССК_M$ остается прежним и равно $(k - 1)$. Число степеней свободы для $ССК_B$ равно $(N - k)$, где $N = n_1 + n_2 + \dots + n_k$. Формулы для сумм квадратов модифицируются:

$$CK_M = \sum_{j=1}^k \frac{\left(\sum_{i=1}^{n_j} x_{ij} \right)^2}{n_j} - \frac{\left(\sum_{j=1}^k \sum_{i=1}^{n_j} x_{ij} \right)^2}{N},$$

$$CK_B = \sum_{j=1}^k \sum_{i=1}^{n_j} x_{ij}^2 - \sum_{j=1}^k \frac{\left(\sum_{i=1}^{n_j} x_{ij} \right)^2}{n_j}.$$

Обсуждаемые здесь проблемы значимости результатов дисперсионного анализа базируются на определенных предположениях, в частности на предположении о нормальности исходных данных и равенстве (однородности) дисперсий. Интерес представляет выяснение того, как нарушения этих предпосылок сказываются на уровне значимости. Теоретические и экспериментальные исследования показали следующее:

1. Когда объемы выборок равны, влиянием неоднородности дисперсий на уровень значимости F -критерия можно пренебречь.

2. Когда объемы выборок и их дисперсии не равны, а из совокупностей с большими дисперсиями выбирается меньшее число объектов, вероятность ошибки первого рода больше.

3. При этих же условиях, если из совокупностей с большими дисперсиями берется большее число объектов, то вероятность ошибки первого рода меньше.

4. Влияние нарушения нормальности на номинальный уровень значимости F -критерия очень незначительно.

Таким образом, процедура дисперсионного анализа достаточно устойчива к нарушению предпосылок, лежащих в ее основе.

8.4. МНОЖЕСТВЕННОЕ СРАВНЕНИЕ СРЕДНИХ

Если ограничиться в результате дисперсионного анализа при непринятии нулевой гипотезы только констатацией факта, что не все средние равны между собой, то это несет не слишком много информации. В самом деле, если при достаточно большом N $\bar{x}_1 = \bar{x}_2 = \bar{x}_3 = 30$, а $\bar{x}_4 = 80$ и $\bar{x}_1 = 10$, $\bar{x}_2 = 30$, $\bar{x}_3 = 50$ и $\bar{x}_4 = 80$, то, очевидно, в обоих случаях однофакторный дисперсионный анализ приведет к отклонению нулевой гипотезы, хотя характер различия между средними совершенно разный. Поэтому результаты дисперсионного анализа по необходимости должны быть дополнены множественным сравнением средних.

Можно попарно сравнить все средние с использованием, например, t -критерия. Однако t -критерий для этих целей непригоден, так как он разрабатывался для сравнения двух выборок, и поэтому рассчитывать на его валидность для определения значимости различий между минимальным и максимальным выборочными средними в совокупности k средних не приходится. В самом деле, если k , например, равно 50 и из одной и той же генеральной совокупности извлекается 50 выборок случайным образом, то t -критерий будет описывать максимальное и минимальное значения как существенно различающиеся гораздо чаще, чем предсказывает выбранный уровень значимости α .

Поэтому были разработаны различные методы множественных сравнений, один из которых рассмотрим, не вдаваясь в его теоретическое обоснование, а обращая внимание на технику вычислений.

Введем понятие студентизированного размаха. Пусть X_1, X_2, \dots, X_k – независимые одинаково распределенные нормальные случайные величины и пусть W_k – размах этой выборки, s^2 – оценка параметра σ^2 , имеющая $n - 1$ степень свободы и не зависящая от W_k . Тогда случайная величина $q_{k,n-1} = W_k / s$ носит название студентизированного размаха.

Для однофакторного анализа студентизированный размах представляет собой разность между максимальным и минимальным средним k независимых выборок по n элементам из нормальной генеральной совокупности, деленную на $\sqrt{ССК_B / n}$, и является критическим значением для выбранного уровня значимости α .

Итак, сначала определяются все $k(k-1)/2$ сравнения между выборочными средними. Например, если $k = 3$, то вычисляются $\bar{x}_1 - \bar{x}_2, \bar{x}_1 - \bar{x}_3, \bar{x}_2 - \bar{x}_3$. Все разности такого вида делятся на $\sqrt{ССК_B / n}$, где n – число наблюдений в группе, соответствующей любому уровню фактора.

Полученные значения сравниваются со студентизированным размахом и делаются соответствующие выводы.

Далее определяются доверительные интервалы для всех разностей средних по формуле

$$(\bar{x}_j - \bar{x}_{j^*}) \pm q_{k,k(n-1)} \Big|_{1-\alpha} \cdot \sqrt{ССК_B / n},$$

где $q_{k,k(n-1)}|_{1-\alpha}$ – критическое значение студентизированного размаха с k и $k(n-1)$ степенями свободы и уровнем значимости α .

ПРИМЕР 8.4. Пусть в результате эксперимента получены значения $\bar{x}_1 = 22,6$; $\bar{x}_2 = 23,4$; $\bar{x}_3 = 28,5$; $n = 11$; $ССК_B = 4,11$.

Имеем $\bar{x}_1 - \bar{x}_2 = -0,8$, $\bar{x}_1 - \bar{x}_3 = -5,9$, $\bar{x}_2 - \bar{x}_3 = -5,1$. Эти разности, деленные на $\sqrt{4,1/11} = 0,61$, дают $-1,31$, $-9,66$, $-8,35$. Из [10] табличное значение критического значения студентизированного размаха со степенями свободы $k = 3$ и $k(n-1) = 30$ и уровнем значимости $\alpha = 0,05$ $q_{3,30} = 3,49$. Следовательно, можно сделать вывод о том, что первое и второе средние незначимо отличаются друг от друга, а третье среднее отличается и от первого, и от второго. Доверительные интервалы для разностей будут выглядеть так:

$$-0,8 \pm 3,49 \cdot 0,61 = (-2,93; 1,33);$$

$$-5,9 \pm 3,49 \cdot 0,61 = (-8,03; -3,77);$$

$$-5,1 \pm 3,49 \cdot 0,61 = (-7,23; -2,97).$$

Дисперсионный анализ и множественное сравнение средних могут быть проведены и для значительно более сложных схем. Последовательность проверки нулевых гипотез, которых может быть достаточно много, сохраняется та же, что и в случае простого однофакторного эксперимента. Сохраняется и способ построения доверительных интервалов.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Дайте наиболее полное определение термина «модель».
2. Каким образом трансформировалось понятие «цель» за весь период развития теории познания?
3. Как Р. Шеннон определяет понятие «имитационное моделирование»?
4. Какие два основных типа тестов применяют для оценки адекватности имитационной модели.
5. В чем отличие теоретико-вероятностной математической модели от статистической?
6. Укажите единицы размерности основных числовых характеристик случайных величин.
7. Можно ли по виду графика определить, является ли функция плотностью распределения случайных величин или функцией распределения?
8. Свяжите причинно-следственными связями понятия независимых случайных величин и некоррелированных случайных величин.
9. Как интерпретируется знак коэффициента корреляции?
10. В чем отличие ковариации от коэффициента корреляции?
11. Чему равна площадь под графиком для функции плотности распределения хи-квадрат?
12. Какой закон называется законом редких явлений?
13. Почему распределение Эрланга считается универсальным?
14. Чем отличается гистограмма частот выборки от функции выборочной плотности?

15. Приведите семейства распределений с числом параметров 1, 2, 3, 4 и более.
16. Перечислите свойства функции распределения и плотности распределения.
17. Сформулируйте алгоритм генерации распределения двумерной случайной величины с помощью функции плотности распределения $f(x, y)$.
18. При проверке критерия согласия для средних по результатам эксперимента делается вывод о том, что почти всегда принимается нулевая гипотеза. Как это характеризует результаты эксперимента?
19. При проверке критерия согласия для дисперсий по результатам эксперимента делается вывод о том, что редко принимается нулевая гипотеза. Как это характеризует результаты эксперимента?
20. Дайте геометрическую интерпретацию двумерной выборки в случае, когда коэффициент корреляции равен ± 1 .
21. Перечислите три гипотезы, которые рекомендуется проверить перед проведением дисперсионного анализа.

ЧАСТЬ 2. ПРИМЕРЫ ВЕРОЯТНОСТНЫХ МОДЕЛЕЙ

9. ВЕРОЯТНОСТНЫЙ КАЛЬКУЛЯТОР ДЛЯ ИНЖЕНЕРНЫХ РАСЧЕТОВ

Разработана иерархия классов на языке C++ для работы с одномерными независимыми случайными величинами. Поддерживаются основные арифметические операции, генерация реализаций. В качестве применения разработано Win-приложение «Вероятностный калькулятор», который вычисляет результат арифметических выражений со случайными величинами и показывает его числовые и функциональные характеристики.

В данной работе описывается библиотека классов для работы с одномерными независимыми случайными величинами. Аналогов подобной библиотеки обнаружено не было. В качестве примера использования данных классов разработано инновационное приложение «Вероятностный калькулятор», которое представляется полезным для научных и инженерных расчетов. В этой связи можно упомянуть моделирование сетевого трафика, интервальные вычисления, прогнозирование транспортных потоков, стохастические финансовые операции и т.п. В подобных задачах довольствуются, как правило, матожиданием и дисперсией результата, представляющего собой случайную величину. Калькулятор способен представить ответ в виде функций плотности и распределения, что гораздо информативнее, чем числовые характеристики.

Приложение «Вероятностный калькулятор» можно использовать как средство для сложения, вычитания, деления, умножения и возведения в степень различных случайных величин из хорошо известных параметрических семейств: равномерное, нормальное, экспоненциальное и др.

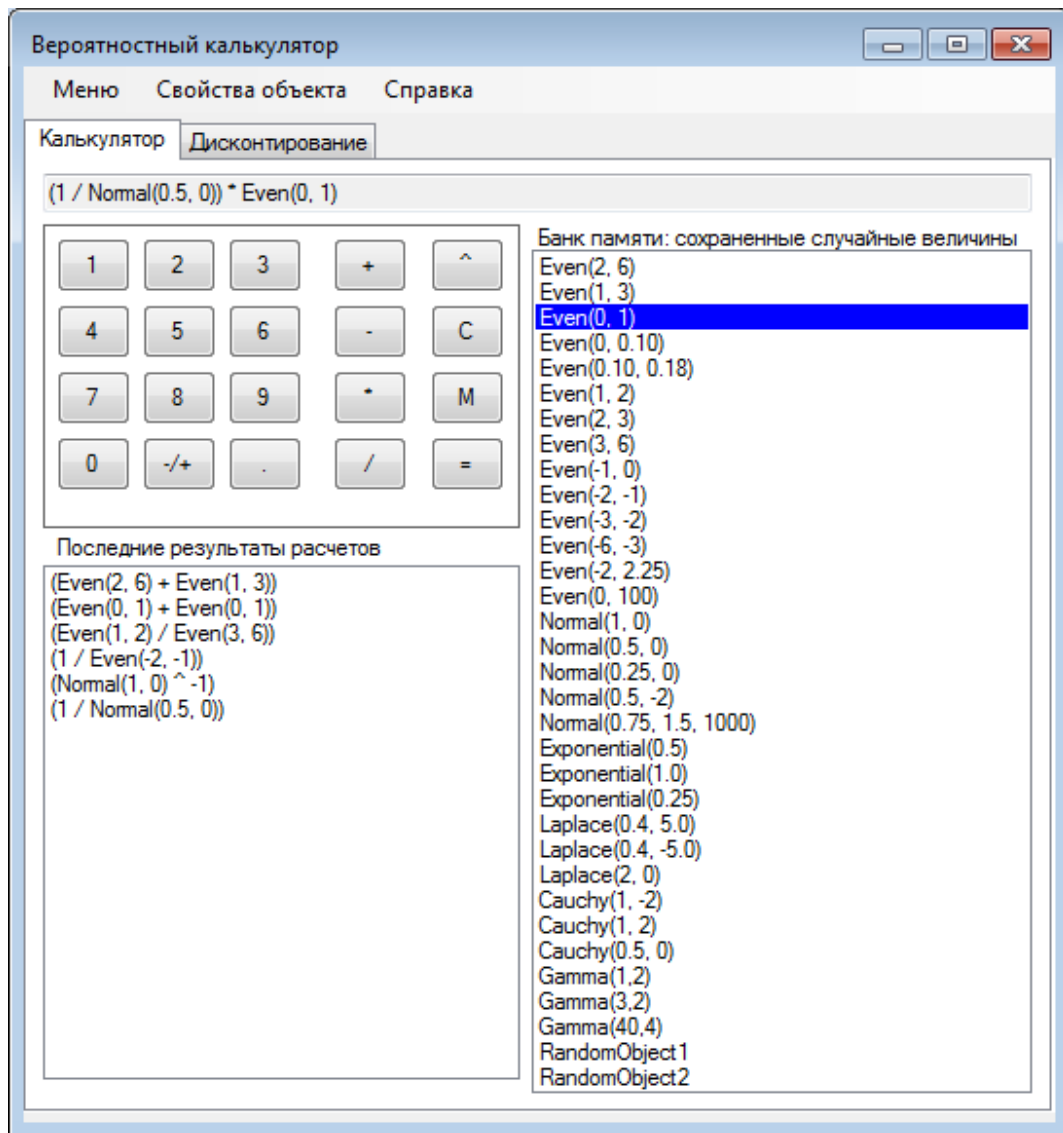
Основная экранная форма приложения представлена на рисунке.

В верхней части формы расположен *редактор*, позволяющий просматривать информацию о текущей операции. Калькулятор работает по принципу одного действия, то есть для выполнения множественных действий следует использовать клавиши памяти «М» и вычисления «=». После получения промежуточного результата с помощью «=» можно продолжать вычисления, выбрав операцию и число/объект.

Слева находится *стандартная панель* калькулятора, позволяющая вводить цифры, выполнять арифметические операции (сложение +, вычитание -, умножение *, деление /, возведение в степень ^), *сохранять* полученный результат в банке памяти (М) и *очищать* редактор (С).

Справа расположен комбинированный список, предназначенный для хранения одномерных случайных величин. В данный список с помощью клавиши «М» после вычисления результата можно помещать полученную новую случайную величину. Для помещения величины из памяти в окно расчетов необходимо щелкнуть по объекту два раза левой клавишей мыши.

Приложение позволяет выводить информацию о случайной величине: числовые характеристики, графики плотности и распределения.



Вероятностный калькулятор

Арифметические операции со случайными величинами реализуются по теоретическим формулам:

– плотность суммы $Z = X + Y$

$$h(z) = \int_{-\infty}^{\infty} f(x) \cdot g(z - x) dx;$$

– плотность разности $Z = X - Y$

$$h(z) = \int_{-\infty}^{\infty} f(x) \cdot g(x - z) dx;$$

– плотность произведения $Z = X \cdot Y$

$$h(z) = \int_{-\infty}^{\infty} f(x) \cdot g\left(\frac{z}{x}\right) \cdot \frac{1}{x} dx;$$

– плотность отношения $Z = X / Y$

$$h(z) = \int_{-\infty}^{\infty} f(x) \cdot g(x / z) \cdot x dx.$$

Рассмотрим новую бинарную операцию $\#$ над случайными величинами X и Y с известными законами распределения. Эта операция позволяет получать малые вариации случайных величин (СВ) в задачах стохастической оптимизации.

Реализации новой случайной величины $Z = X \# Y$ строятся по правилу чередования: четные номера берутся в соответствии с X , нечетные – в соответствии с Y . Найдем закон распределения Z .

Методом «обратной функции» можно генерировать реализации случайной величины X с законом распределения F : $x = F^{-1}(r)$, где r – равномерно распределенная случайная величина на отрезке $[0, 1]$.

Пусть $f(x)$, $g(y)$ и $h(z)$ – функции плотностей распределения случайных величин X , Y и Z соответственно. Функции f и g известны, найдем h .

Оценим вероятность попадания Z в маленький интервал dz . По формуле полной вероятности с учетом чередования получаем

$$\begin{aligned} P(z \leq Z \leq z+dz) &= h(z)dz = 1/2 \cdot P(z \leq X \leq z+dz) + 1/2 \cdot P(z \leq Y \leq z+dz) = \\ &= 1/2 \cdot f(z)dz + 1/2 \cdot g(z)dz. \end{aligned}$$

Отсюда после сокращения на произвольное приращение dz получаем ответ

$$h(z) = 1/2 \cdot f(z) + 1/2 \cdot g(z).$$

Задача обобщается следующим образом. Пусть имеется однопараметрическое семейство случайных величин $V(\alpha)$ с плотностями распределения $f(x, \alpha)$. Выберем параметр α_0 в соответствии с известной плотностью $g(\alpha)$ и в качестве реализации неизвестной случайной величины Z возьмем реализацию известной случайной величины $V(\alpha_0)$.

Тогда плотность распределения Z имеет вид

$$h(z) = \int_{-\infty}^{+\infty} g(\alpha) \cdot f(z, \alpha) d\alpha.$$

Основной целью данной работы явилось создание иерархии классов. Классы принадлежат одному пространству имен. Это объединяет разработанные классы в иерархию классов, производных от одного базового абстрактного класса с поддержкой механизма виртуальных функций.

В основе иерархии лежит класс *Function*, содержащий общие методы для двух производных классов: *Density* и *Distribution*. Эти два класса предназначены для работы соответственно с функциями плотности и с функциями распределения. Путем множественного наследования из классов *Density*, *Distribution* и класса генерации реализаций *IGenerate* произведен класс *RandomObject*. Этот класс вычисляет числовые характеристики случайной величины, перегружает арифметические операции. Из класса *RandomObject* наследуются основные параметрические семейства случайных величин. Программист может создать собственное семейство путем стандартного наследования и переопределения соответствующих методов.

Библиотека классов содержит конструкторы, которые позволяют создавать объекты классов на основе функции распределения или функции плотности распределения. Эти функции могут быть заданы программным способом, таблицей в файле или массивом в оперативной памяти.

Для конкретных объектов случайных величин вычисляются их основные числовые характеристики: математическое ожидание, мода, медиана, дисперсия, среднеквадратичное отклонение, коэффициент асимметрии и эксцесс.

10. АНАЛИЗ ОБМЕННЫХ КУРСОВ ВАЛЮТ В БАНКАХ ГОРОДА

На многих отечественных сайтах можно найти информацию об обменных курсах основных мировых валют: американский доллар и евро. Например, на сайте *e1.ru* можно найти информацию об обменных курсах более чем в 80 обменных пунктах и банках г. Екатеринбурга. При этом данные, сведенные в таблицу, подвергаются минимальной статистической обработке. Цветом выделяются минимальные и максимальные курсы. Указываются средневзвешенные курсы и курсы Центробанка.

В то же время данные хранят важную информацию, для понимания которой надо иметь минимальные сведения из теории математической статистики. Можно предположить, что уровень математической и экономической грамотности населения позволяет воспринять эту информацию и применять ее для своей пользы.

Скопируем данные со страницы сайта от 19 апреля 2011 года в приложение *MS Excel*. Проведем предварительную ручную обработку данных. Наименования банков можно удалить – нам достаточно иметь обезличенные данные. Проведем массовую замену разделителя десятичных знаков с точки на запятую. Заменяем появившиеся кое-где даты типа «28.дек» на число 28,12.

Итак, в таблице есть колонки: курсы продажи/покупки американского доллара и курсы продажи/покупки евро. При этом курсы указываются с точки зрения банков. Это банки продают и покупают валюту по данным курсам. Построим гистограммы каждого из 4 курсов.

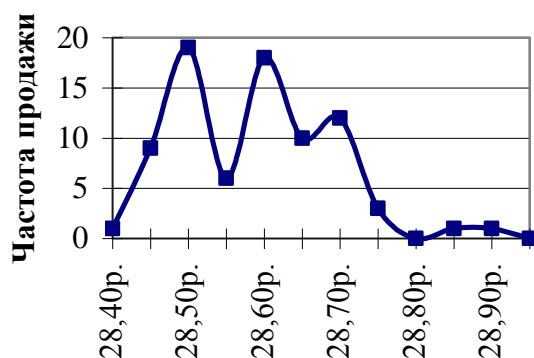
Каждую колонку будем обрабатывать по отдельности. Отсортируем колонки по возрастанию. Создадим диапазон карманов, начиная с минимального числа в колонке и заканчивается максимальным. Количество карманов возьмем, например, 10. Для построения гистограмм (см. рисунок), воспользуемся сервисной функцией «Сервис – Анализ данных – Гистограмма».

Рассчитаем коэффициенты асимметрии и эксцесса и проанализируем ситуацию (см. таблицу).

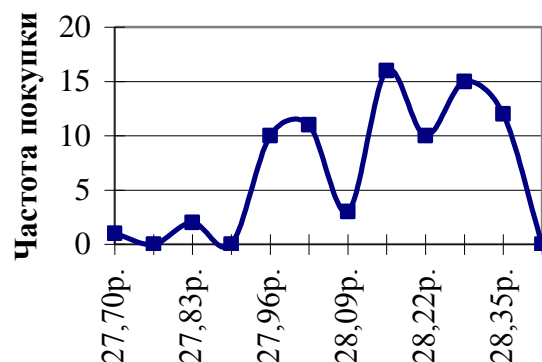
Расчетные коэффициенты анализа обменных курсов валют

Коэффициенты анализа	Американский доллар		Евро	
	продажа	покупка	продажа	покупка
Асимметрия	0,54	-0,50	2,55	-1,57
Эксцесс	0,16	-0,43	8,16	5,34
Степень желанья банков совершать данную операцию	Хотят	Хотят	Хотят больше	Хотят больше
Наличие доминирующих факторов влияния на обменный курс	Нет	Нет	Есть	Есть

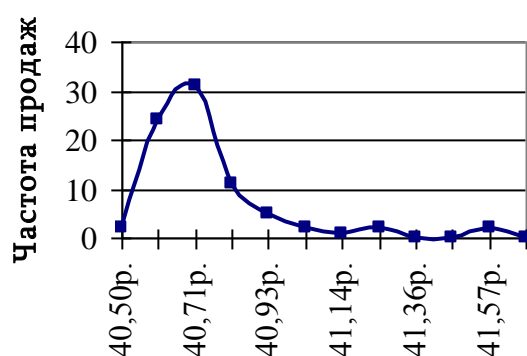
а



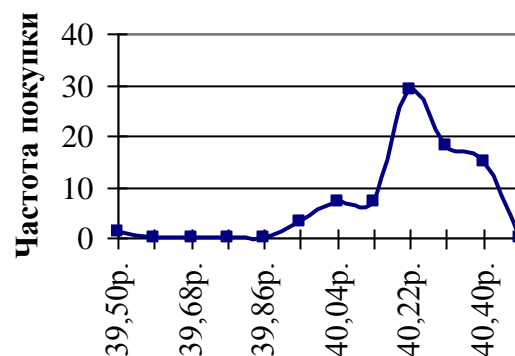
б



в



г



Гистограммы курсов валют:

а – продажа доллара; б – покупка доллара; в – продажа евро; г – покупка евро

Асимметрия говорит о совокупном желании/нежелании банков продавать или покупать валюту.

При этом положительная асимметрия при продаже означает, что у банков имеется желание продавать валюту. Аналогично, отрицательная асимметрия при покупке означает, что у банков имеется желание покупать валюту.

И наоборот, отрицательная асимметрия при продаже означает, что гистограмма скошена влево, то есть большая часть банков предлагает к продаже валюту по заниженной цене, поэтому у банков имеется желание продавать валюту.

Практический вывод. Допустим, банки проявили сильное желание покупать доллары (вместо -0,50 мы, например, имеем -10). По-видимому, следует доверять банковскому сообществу, поэтому физическим лицам следует делать то же самое.

Величина эксцесса показывает отклонение гистограммы от кривой нормального распределения. Эксцесс нормальной кривой равен нулю. При отрицательном эксцессе мы имеем туповершинные кривые, при положительном эксцессе кривые островершинные. Отклонение гистограммы от нормальной кривой свидетельствует о возможном наличии одного или малого числа скрытых факторов, влияющих на валютный рынок. В предкризисных ситуациях асимметрия и эксцесс увеличиваются.

Ситуация на валютном рынке быстро меняется, поэтому анализ курсов следует проводить в динамике. Важно знать, падает или возрастает желание банков совершать валютные операции.

Таким образом, полученные статистические характеристики обменных курсов могут дополнить стандартные страницы сайтов и при наличии соответствующих комментариев будут полезны населению.

11. МОДЕЛИРОВАНИЕ ИГРОВОГО АВТОМАТА

В 1884 году американец Чарльз Фай собрал в своей автомастерской первую слот-машину (автомат или аппарат), дав ей громкое название «Колокол свободы» («Liberty Bell»). На смену слот-автоматам в 1891 году пришли покерные аппараты. Их авторство тоже принадлежит Чарльзу Фаю.

Что касается видеослотов, всевозможных онлайн-казино, онлайн-покерных столов, рулеток, то их история началась примерно в одно время с историей самого Интернета. Первое Интернет-казино появилось в 1994 году.

Задача моделирования игрового автомата аналогична, например, техническим задачам оценки надежности оборудования. Рассмотрим пример. Имеется прибор с определенными вероятностями различных неисправностей. Каждая неисправность требует некоторой фиксированной стоимости ремонта. При покупке прибора пользователь может получить как значительные выгоды от его использования, так и затраты на его несвоевременный ремонт. Возникает задача: оценить риск приобретения данного прибора, сравнить аналогичные приборы различных марок, предложить варианты модификаций прибора с целью повышения его качества и уменьшения риска денежных затрат на его ремонт.

В основе модели игрового автомата лежит некоторая вероятностная задача. СВ X этой задачи может иметь различную природу. Она может быть одномерная или многомерная, дискретная или непрерывная. X может задаваться аналитически с помощью законов распределения или экспериментально, путем создания генератора или алгоритма.

Рассмотрим одномерную дискретную СВ с заданным законом распределения (табл. 11.1). Если игровой автомат построен на базе непрерывной СВ, то табл. 11.1 представляет собой функцию распределения. Модификация этой функции может осуществляться с помощью арифметических операций, реализованных в библиотеке иерархии классов СВ, или с помощью вероятностного калькулятора (см. главу 9).

Таблица 11.1

Закон распределения X

Исход X	x_1	x_2	...	x_n
P	p_1	p_2	...	p_n

Очевидно, $\sum_{i=1}^n p_i = 1$, что является необходимым условием правильности составления табл. 11.1.

На основании этой задачи построим соотношения для игрового автомата.

Введем обозначения:

g – стоимость жетона для одной игры, руб.;

v_i – цена выигрыша игрока для исхода x_i , руб., $v_i \geq 0, i = 1, \dots, n$;

t_0 – продолжительность одной игры, ч;

N – количество игр в 1 час, $N = \frac{1}{t_0}$;

T – желаемое время нахождения игрока в казино, ч;

r – доходность игрового автомата по простой процентной ставке в день, %;

S – наличная денежная сумма игрока, руб.;

Q – количество часов, которые игрок фактически провел в казино, ч.

Обозначим через V случайную величину выигрыша игрока за одну игру, без учета стоимости жетона (табл. 11.2). Распределение величины V задается заранее и не изменяется в процессе игры.

Таблица 11.2

Закон распределения V выигрыша игрока

V	v_1	v_2	\dots	v_n
P	p_1	p_2	\dots	p_n

Средний выигрыш игрока за одну игру без учета стоимости жетона

$$M(V) = \sum_{i=1}^n p_i v_i.$$

Дисперсия выигрыша V равна

$$D(V) = \sum_{i=1}^n (v_i - M(V))^2 p_i.$$

Одна игра представляет собой стохастическую финансовую операцию, для которой начальная сумма известна и равна стоимости жетона g ; конечная сумма – случайная величина V ; срок операции t_0 . Рассчитаем доходность r этой операции по формуле простой процентной ставки за 1 час:

$$r = \frac{V - g}{gt_0}.$$

Почасовая доходность r является случайной величиной с числовыми характеристиками

$$M(r) = \frac{M(V) - g}{gt_0}, \quad (11.1)$$

$$\sigma(r) = \frac{1}{gt_0} \sigma(V). \quad (11.2)$$

Величина $\sigma(r)$ является мерой риска стохастической финансовой операции.

Если риск $\sigma(r)$ равен 0, то игра превращается в детерминированную операцию обмена жетона g на некоторую меньшую фиксированную сумму, например, 4 руб. 90 коп. В результате такой игры игрок имеет гарантированный проигрыш 10 коп. Доходность операции для игрока равна

$$(4,9 - 5)/(5t_0) = -0,02/t_0 = -2/t_0, \% \text{ в час.}$$

А игрового азарта нет!

Другой важной характеристикой случайной величины является коэффициент асимметрии

$$A(V) = \frac{1}{\sigma^3(V)} \sum_{i=1}^n (v_i - M(V))^3 p_i.$$

Если коэффициент асимметрии меньше 0, то плотность распределения смещена влево, иначе – вправо. Это свойство отражается на игровом автомате следующим образом: при положительном скосе игрок часто выигрывает маленькие суммы и редко проигрывает большие. При отрицательном скосе – наоборот, игрок часто проигрывает малые суммы и изредка получает крупные. При этом можно изменять скос, сохраняя математическое ожидание и дисперсию.

Если риск $\sigma(r)$ финансовой операции со случайной доходностью r больше средней доходности $M(r)$, то такая операция считается высокорискованной. В нашем случае игра для клиента заведомо высокорискованная, так как $M(r) < 0$, а $\sigma(r) > 0$ и, следовательно, $\sigma(r) > M(r)$.

Для казино ситуация другая. С каждой игры казино получает в среднем доход $-M(r)$, и доходность этой операции равна $-M(r) > 0$. Для казино риск не наступает, когда

$$-M(r) > \sigma(r). \quad (11.3)$$

Далее, казино заинтересовано в том, чтобы игрок, в среднем, просидел в заведении T часов. За одну игру он теряет в среднем сумму $M(V) - g$. За время T игрок сыграет NT игр и в среднем проиграет сумму, равную $(M(V) - g)NT$. Это – при условии возможности неограниченного займа денег. В идеальном для игрока и казино случае игрок за время T проиграет всю запланированную для игрового дня сумму. Поэтому

$$-(M(V) - g)NT = S. \quad (11.4)$$

Обозначим Z – фактический проигрыш игрока (с учетом стоимости жетонов) за Q часов, то есть за QN независимых игр. Таким образом,

$$Z = \sum_{i=1}^{QN} (V - g),$$

то есть Z – сумма большого числа независимых случайных величин, ни одно из которых не оказывает решающего влияния на всю сумму. По теореме Ляпунова Z – почти нормальная СВ.

Найдем числовые характеристики Z .

$$M(Z) = QN(M(V) - g), \quad (11.5)$$

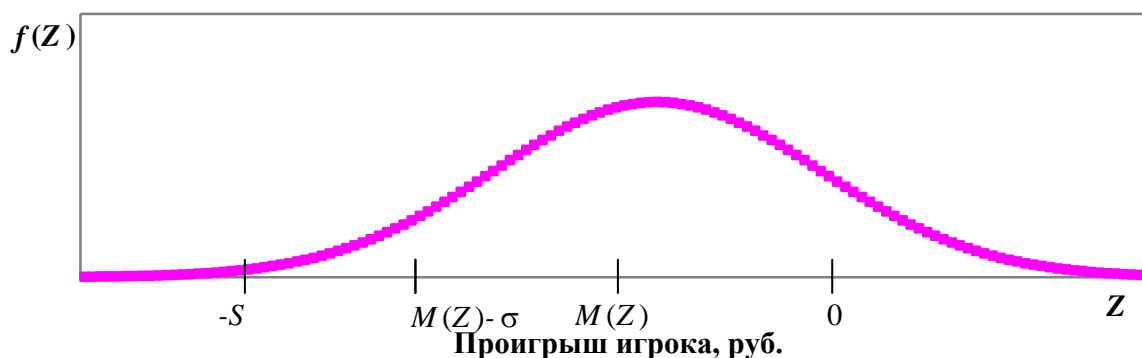
$$\sigma(Z) = QN \cdot \sigma(V). \quad (11.6)$$

Плотность распределения Z представлена на рисунке.

Игрок имеет в наличии S рублей. Допустим, он не может взять денег в долг. Тогда если $Z < -S + g$, то игрок не сможет купить новый жетон и покидает казино. Найдем вероятность того, что игрок покинет казино за Q часов:

$$P(Z < -S + g) = F_0\left(\frac{-S + g - M(Z)}{\sigma(Z)}\right), \quad (11.7)$$

где F_0 – функция распределения нормированной нормальной СВ.



Плотность распределения проигрыша игрока за Q часов

Зададимся некоторой критической вероятностью $P_{\text{крит}}$ ухода игрока через Q часов после прихода в казино. Подставляя в (11.7) выражения (11.5) и (11.6), получаем соотношение

$$F_0\left(\frac{-S + g - QN(M(V) - g)}{QN \cdot \sigma(V)}\right) = P_{\text{крит}}, \quad (11.8)$$

Итак, основные характеристики игрового автомата определяются следующими соотношениями:

1. Выражение (11.1) определяет среднюю почасовую доходность игрового автомата по простой процентной ставке. Для перехода к годовой простой процентной ставке надо r умножить на коэффициент $24 \cdot 365 = 8760$. Эта ставка ограничивается сверху законодательным путем⁵ $r < r_{\text{макс}}$. Если $r_{\text{макс}}$ не очень велико, то игровой автомат перестает быть «одноруким» бандитом.

2. Величина (11.2) отражает степень риска игрового автомата для игрока.

3. Неравенство (11.3) определяет, является ли игровой автомат высокорискованным для казино. Для игрока игровой автомат – заведомо высокорискованный.

4. Уравнение (11.4) определяет среднее время T пребывания в казино игрока, имеющего в наличии S денег и возможность неограниченного займа.

5. Соотношение (11.8) определяет долю $P_{\text{крит}}$ игроков, которые имеют в наличии S денег без возможности неограниченного займа и покинули казино через Q часов. Для казино желательно, чтобы $Q \approx T$.

6. В процессе игры закон распределения выигрышей вряд ли можно пересмотреть. Но вероятности p_i из табл. 11.1 зависят от параметров вероятностной задачи. Изменение этих параметров может привести к плавному изменению характеристик в нужную сторону. Например, плавное уменьшение риска (12.2) может уменьшить болезненные симптомы игроманов.

ПРИМЕР 11.1. В урне I находится 5 белых и 7 черных шаров, а в урне II – 3 белых и 5 черных шаров. Из I урны во II перекладываются 3 шара, затем из II урны вынимаются 3 шара. X – количество белых шаров, вынутых из II урны.

⁵ Федеральный закон от 29 декабря 2006 года № 224-ФЗ «О государственном регулировании деятельности по организации и проведению азартных игр и о внесении изменений в некоторые законодательные акты Российской Федерации».

Найдем закон распределения X . Сформулируем гипотезы:

H_0 – из первой урны во вторую переложено 0 белых и 3 черных шара.

H_1 – из первой урны во вторую переложено 1 белый и 2 черных шара.

H_2 – из первой урны во вторую переложено 2 белых и 1 черный шар.

H_3 – из первой урны во вторую переложено 3 белых и 0 черных шар.

Вероятности гипотез найдем по классической формуле теории вероятностей $P(A) = m/n$, где m – число исходов благоприятных событию A ; n – общее число исходов. При этом считаем размещения, т. е. предполагаем, что все шары пронумерованы, и порядок вынимания шаров значим.

$$P(H_0) = 7 \cdot 6 \cdot 5 / (12 \cdot 11 \cdot 10) = 15,91 \%,$$

$$P(H_1) = 5 \cdot 7 \cdot 6 / (12 \cdot 11 \cdot 10) = 48,73 \%,$$

$$P(H_2) = 5 \cdot 4 \cdot 7 / (12 \cdot 11 \cdot 10) = 31,82 \%,$$

$$P(H_3) = 5 \cdot 4 \cdot 3 / (12 \cdot 11 \cdot 10) = 4,55 \%,$$

Найдем условные вероятности события $X = 0$ при этих гипотезах:

$$P(X = 0 | H_0) = 8 \cdot 7 \cdot 6 / (11 \cdot 10 \cdot 9) = 33,94 \%,$$

$$P(X = 0 | H_1) = 7 \cdot 6 \cdot 5 / (11 \cdot 10 \cdot 9) = 21,21 \%,$$

$$P(X = 0 | H_2) = 6 \cdot 5 \cdot 4 / (11 \cdot 10 \cdot 9) = 12,12 \%,$$

$$P(X = 0 | H_3) = 5 \cdot 4 \cdot 3 / (11 \cdot 10 \cdot 9) = 6,06 \%,$$

По формуле полной вероятности получаем

$$p_0 = P(X = 0) = P(H_0) \cdot P(X = 0 | H_0) + \dots + P(H_3) \cdot P(X = 0 | H_3) = 19,66 \%,$$

Аналогично находим $p_1 = 48,51 \%$, $p_2 = 28,10 \%$, $p_3 = 3,73 \%$.

Результаты расчетов запишем в табл. 11.3.

Таблица 11.3

Закон распределения X

Исход X	0	1	2	3
P , %	19,66	48,51	28,10	3,73

Размер выигрышей выберем произвольно (табл. 11.4).

Таблица 11.4

Таблица выигрышей

Исход X	0	1	2	3
Выигрыш V , руб.	0,00	3,90	4,00	20,00

Объединяя табл. 11.3 и 11.4, получаем табл. 11.5.

Таблица 11.5

Закон распределения выигрышей

Выигрыш V , руб.	0,00	3,90	4,00	20,00
P , %	19,66	48,51	28,10	3,73

Таким образом, $M(V) = 3,76$ руб.

Зададим параметры модели (условия игры):

- стоимость жетона $g = 5$ руб.;
- продолжительность одной игры $t_0 = 0,10$ ч = 6 мин;
- количество игр за 1 час $N = 10,0$;
- желаемое время нахождения игрока в казино $T = 5,0$ ч;

- наличная денежная сумма игрока $S = 100,00$ руб.;
- время, которое игрок фактически проводит в казино, $Q = 5,0$ ч.

Результаты анализа модели.

1. Средняя почасовая доходность игрового автомата (простая процентная ставка) $M(r) = -247,50 \%$.

СКО почасовой доходности $\sigma(r) = 711,32 \%$.

Таким образом, игровой автомат для казино является высокорискованным.

2. Средний проигрыш за время T при неограниченном бюджете равен $61,87 < S = 100$ руб.

Поэтому игровому заведению выгодно, например, уменьшить средний выигрыш игрока с величины 3,76 руб. (см. табл. 11.5) до величины

$$M(V) = g - S/(NT) = 5 - 100 / (10 \cdot 5) = 3,00 \text{ руб.}$$

3. Среднее время, за которое игрок потратит заданную сумму S при неограниченном займе равно

$$T' = -S/([M(V) - g] \cdot N) = 8,1 \text{ ч} > T.$$

Поэтому казино заинтересовано в уменьшении величины T , например, путем увеличения N , т.е. уменьшая время игры t_0 .

4. Найдем вероятность $P(Z < -S + g)$ ухода игрока из казино за время Q .

$$P(Z < -S + g) = 42,61 \%.$$

Другими словами, 42,61 % игроков, имеющих сумму $S = 100$ руб., покинут казино через $Q = 5,0$ часов.

Предложенные соотношения позволяют модифицировать модель игрового автомата в 3 этапа.

1-й этап. Найти траекторию и динамику изменения параметров модели игрового автомата (доходность, риск).

2-й этап. Определение необходимых для этого законов распределения (вероятности p_i из табл. 11.1).

3-й этап. Поиск параметров вероятностной задачи, при которой имеют место данные вероятности. В приведенном примере это количество шаров в урнах, количество самих урн и т.д.

Таким образом, модель игрового автомата позволяет управлять доходностью и риском автомата. Плавное, незаметное для игрока, уменьшение риска может уменьшить игровую зависимость. Малая доходность автомата со значительным риском превращает игру в малозатратное времяпрепровождение.

12. МНОГОПАРАМЕТРИЧЕСКАЯ СТОХАСТИЧЕСКАЯ МОДЕЛЬ «УЗОРЫ»

12.1. ИЗОБРАЖЕНИЯ С КОНЕЧНЫМ НАБОРОМ ДАННЫХ

Большинство математических моделей сложных систем определяются фиксированным набором данных. Эти данные могут иметь различную природу. Они могут быть детерминированными или стохастическими, могут иметь разную размерность. Данные могут задаваться функционально или конечным множеством чисел, символов, правил и т.д.

Модель может по-разному реагировать на изменение своих данных.

Устойчивость моделей к малым возмущениям своих параметров свойственна многим техническим системам и математическим конструкциям. В этом случае актуальной является оценка чувствительности модели к возмущениям, возможна линеаризация связей, применение широкого класса методов оптимизации.

Модели с неустойчивым поведением требуют применения робастных подходов, методов регуляризации, масштабирования.

Результаты функционирования модели могут иметь геометрическую интерпретацию. Иногда, построение изображений является основной задачей модели. В этом случае модель настраивается таким образом, чтобы ее графический образ имел заданные характеристики. При всей сложности геометрических образов модели, как правило, возможностей параметров модели недостаточно для установления однозначного соответствия между моделями и реальностью предметной области.

Рассмотрим несколько примеров построения изображений, которые в конечном счете определяются конечным набором данных целого и вещественного типов.

12.2. УЗОРЫ

В литературе сложно найти четкое определение понятия «узор». Мы будем понимать под узором некое целое – изображение, текст, последовательность событий, что угодно, составленное из элементов, между которыми невооруженным глазом видна зависимость и которые являются частью некоторой последовательности. Для описания этой зависимости используются различные математические модели с широким спектром математических методов.

Одной из наиболее известных математических моделей узоров являются фракталы. Понятия фрактал и фрактальная геометрия, появившиеся в конце 1970-х годов прочно вошли в обиход математиков и программистов. Слово фрактал образовано от латинского *fractus* и в переводе означает состоящий из фрагментов. Оно было предложено Бенуа Мандельбротом в 1975 году для обозначения нерегулярных, но самоподобных структур, которыми он занимался. Рождение фрактальной геометрии принято связывать с выходом в 1975 году книги Мандельброта [15]. В его работах использованы научные результаты других ученых, работавших в период 1875–1925 годов в той же области (Пуанкаре, Фату, Жюлиа, Кантор, Хаусдорф). Но только в наше время удалось объединить их работы в единую систему.

Одним из основных свойств фракталов является само подобие. В самом простом случае небольшая часть фрактала содержит информацию о всем фрактале.

Определение фрактала, данное Мандельбротом, звучит так: «Фракталом называется структура, состоящая из частей, которые в каком-то смысле подобны целому».

Одно из главных направлений в использовании фракталов – это машинная графика. С помощью них можно создать или описать поверхности очень

сложной формы. Изменение всего нескольких коэффициентов в уравнении модели позволяет добиться практически бесконечных вариантов исходного изображения. Фрактальная геометрия незаменима при генерации искусственных облаков, гор, поверхности моря. Фактически найден способ легкого представления сложных неевклидовых объектов, образы которых весьма похожи на природные.

По форме представления фракталы подразделяются на: геометрические, алгебраические, аналитические, стохастические, линейчатые.

12.2.1. Геометрические фракталы

Геометрические фракталы являются наиболее наглядными. В двумерном случае их получают с помощью некоторой ломаной (или поверхности в трехмерном случае), называемой генератором. За один шаг алгоритма каждый из отрезков, составляющих ломаную, заменяется на ломаную-генератор в соответствующем масштабе. В результате бесконечного повторения этой процедуры получается геометрический фрактал. Классическим примером такого фрактала является снежинка Коха (рис. 12.1). Граница снежинки, придуманной Хельгом фон Кохом в 1904 году [16], описывается кривой, составленной из трех одинаковых фракталов.

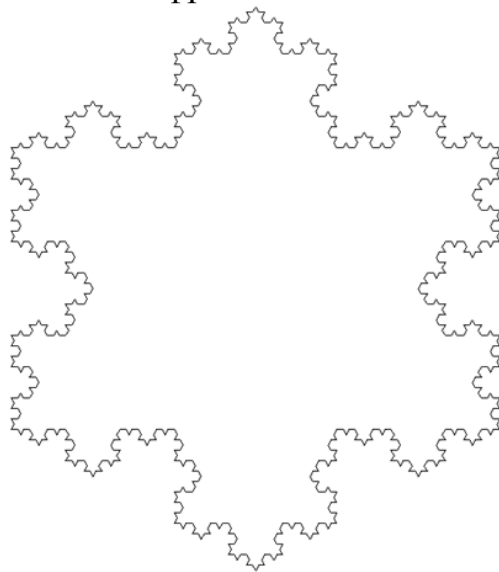


Рис. 12.1. Снежинка Коха

Основой снежинки являются стороны равностороннего треугольника. На базе средней трети стороны построим снова равносторонний треугольник и уберем основание нового треугольника, как показано на рис. 12.2

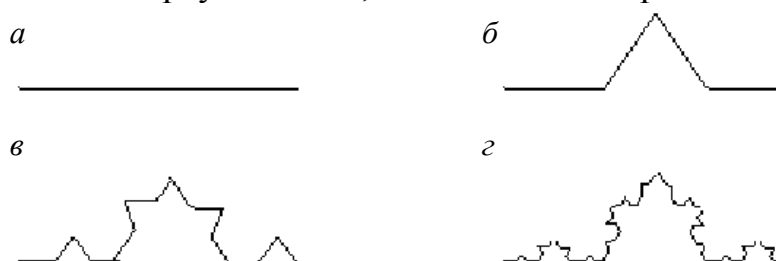


Рис. 12.2. Процесс построения снежинки Коха

Повторяя данную процедуру многократно, получим кривую произвольно большой длины, ограничивающую конечную площадь.

Поясним, что такое фрактальная размерность. Предположим, что фигура заключена в некоторый квадрат. Закроем эту фигуру меньшими квадратами, подобными исходному с коэффициентом $k \gg 1$. Допустим, для этой цели нам хватило $N(k)$ меньших квадратов. Тогда фрактальная размерность равна

$$D = \lim_{k \rightarrow \infty} \ln(N(k)) / \ln(k).$$

Снежинка Коха имеет дробную фрактальную размерность $D = \ln 4 / \ln 3 \approx 1,2618$, которая меньше ее топологической размерности, равной 2.

12.2.2. L-системы

Одним из способов математического описания геометрических фракталов являются *L*-системы.

В основе *L*-систем лежит постоянная перезапись элементов некоторой простой начальной строки. По некоторым правилам производятся замены подстрок этой строки на другие подстроки. В результате визуализации длинной итоговой строки получаем сложные геометрические объекты. Подобный алгоритм был впервые предложен в 1904 году математиком Хельгом фон Кохом для генерации кривой, огибающей снежинку. В шестидесятых годах Хомский [17] использовал эти алгоритмы для описания формальных грамматик. А в 1968 году они были применены венгерским биологом и ботаником Аристидом Линденмайером [18], по имени которого и названы *L*-системы, для изучения развития простых многоклеточных организмов. Позднее область применения *L*-систем была расширена, и они использовались для моделирования сложных ветвящихся структур древовидного типа.

Формально *L*-система состоит из алфавита, начального кодового слова и набора порождающих правил. Правила указывают, как следует преобразовывать кодовое слово при переходе от одной итерации к другой.

Допустим, алфавит состоит из символов $\{F, +, -\}$. В качестве порождающего правила возьмем замену $F = F-F++F-F$. Обновление символов F в данном слове предполагается одновременным, т. е. символы одной итерации обновляются раньше любой буквы следующей итерации.

Для графической реализации *L*-систем в качестве подсистемы вывода обычно используется так называемая «черепашья» графика. При этом точка («черепашка») движется по экрану дискретными шагами, как правило, прочерчивая свой след, но при необходимости может перемещаться без рисования. «Черепашка» обучена интерпретировать и выполнять последовательность команд, задаваемых кодовым словом, буквы которого читаются слева направо. Кодовое слово представляет собой результат работы *L*-системы и может включать только символы из алфавита. Часть символов алфавита реализуют predetermined действия.

Например, в таблице описаны действия 5 символов алфавита:

Алгоритмические действия символов алфавита

Символ алфавита	Действие
[Извлечь положение координат пера из стека
]	Запомнить положение координат пера в стеке
+	Увеличить угол перемещения a на величину α
-	Уменьшить угол a на величину α
F	Переместить перо на шаг h под углом a , прорисовывая след
B	Переместить перо на шаг h под углом a , не прорисовывая след

Размер шага h , начальный угол a и величина α приращения угла a задаются заранее и остаются неизменными для всех перемещений «черепашки». Если начальное направление движения a (угол, отсчитываемый от положительного направления оси Ox) не указано, то полагаем a равным нулю.

Символы +, -, [и] играют вспомогательную роль и не имеют графического образа. Они не обновляются, а просто остаются на тех местах, где встретились.

ПРИМЕР 12.1. L -система, соответствующая снежинке Коха, задается следующим образом:

$$a = 0, \alpha = \pi/3, h = 1.$$

Начальное слово: $F++F++F$.

Порождающее правило: $F = F-F++F-F$.

Графическое представление аксиомы – равносторонний треугольник. В самом деле, «черепашка» делает один шаг вперед, затем угол увеличивается на $2\pi/3$ и «черепашка» делает второй шаг. Наконец, угол увеличивается еще на $2\pi/3$ и «черепашка» делает третий шаг.

На первом шаге каждая буква F в аксиоме $F++F++F$ заменяется на $F-F++F-F$. В результате первого шага получаем слово:

$$(F-F++F-F)++(F-F++F-F)++(F-F++F-F).$$

Повторяя этот процесс, на втором шаге получим слово:

$$F-F++F-F-F-F++F-F++F-F++F-F-F-F++F-F+F-F++F-F-F-F++F-F++F- \\ F++F-F-F-F++F-F+ F-F++F-F- F-F++F-F++F-F++F-F-F-F++F-F$$

и так далее.

Мозаики Пенроуза, которые представляют собой нетранзитивные узоры с группой поворотов 5-го порядка (рис. 12.3), построены с помощью приложения.⁶

⁶ На кафедре автоматики и информационных технологий УрФУ (ait.ustu.ru) под руководством С.П. Трофимова в ходе выполнения дипломной работы «Программа для моделирования и визуализации развития объектов с помощью L -систем» (2009) студентом А.Б. Грабежевым разработано приложение для построения произвольных L -систем.

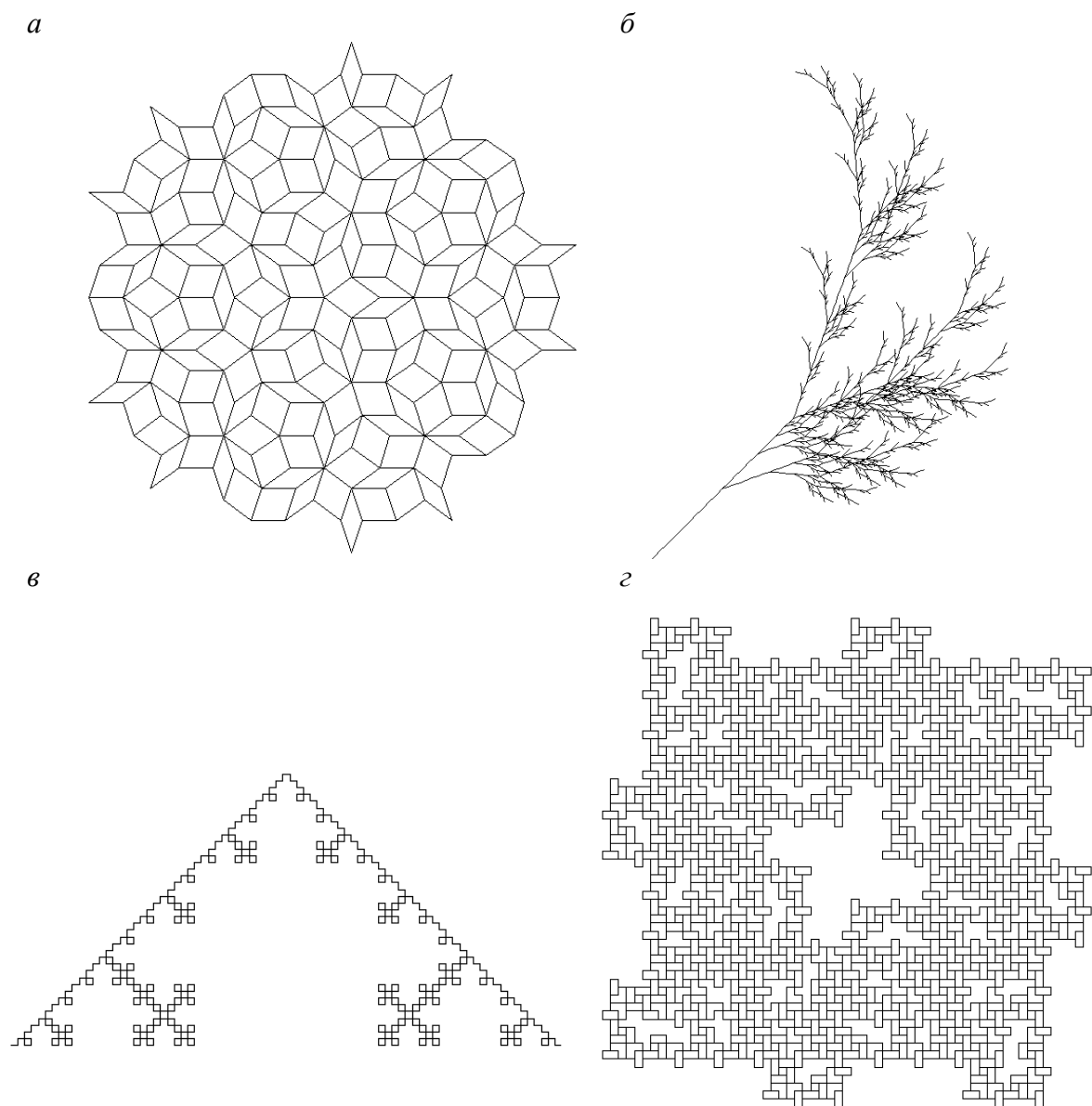


Рис. 12.3. Узоры, построенные с помощью L -систем в приложении [5]:
 а – мозаика Пенроуза; б – фрактальное дерево; в – кривая Коха; г – паркет

12.2.3. Алгебраические фракталы

Алгебраические фракталы получают с помощью нелинейных процессов в n -мерных пространствах. Наиболее изучены двумерные процессы. Представляя нелинейный итерационный процесс как результат расчета дискретной динамической системы, можно воспользоваться терминологией теории этих систем: фазовый портрет, установившийся процесс, аттрактор и т.д.

Известно, что нелинейные динамические системы обладают несколькими устойчивыми состояниями. То состояние, в котором оказалась динамическая система после некоторого числа итераций, зависит от ее начального состояния. Поэтому каждое устойчивое состояние – аттрактор – обладает некоторой областью начальных состояний, из которых система обязательно попадет в рассматриваемые конечные состояния. Таким образом, фазовое пространство системы разбивается на области притяжения аттракторов. Если фазовое пространство является двумерным, то, окрашивая области притяжения к

различным аттракторам различными цветами, можно получить цветовой фазовый портрет этой системы (итерационного процесса). Меняя алгоритм выбора цвета, можно получить сложные фрактальные картины с причудливыми многоцветными узорами. С помощью простых алгоритмов создавать очень сложные нетривиальные структуры. Самым известным примером алгебраического фрактала является множество Мандельброта [15]. Оно получается как множество точек z на комплексной плоскости, для которых итеративная последовательность

$$z_0 = 0, \\ z_{n+1} = z_n^2 + c.$$

не уходит на бесконечность. Модификации множества Мандельброта задаются ограниченным набором параметров, изменение которых влечет значительные изменения в узорах.

12.2.4. Стохастические фракталы

Моделирование с помощью фракталов, которые являются статическими фигурами, вполне приемлемо до тех пор, пока не возникает необходимость рассмотрения таких динамических природных явлений, как падающие потоки воды, турбулентные завихрения дыма, метеосистемы и потоки на выходе реактивных двигателей. В этих случаях один-единственный фрактал соответствует моментальному снимку данного феномена. Если в итерационном процессе случайным образом менять какие-либо его параметры, то узоры начинают динамически изменяться. Структуры, изменяющиеся во времени, моделируются динамическими системами.

Фракталы позволяют описывать регулярные структуры, в основе которых лежит самоподобие и рекурсия. Естественные узоры обычно образуются в ходе вырастания одного элемента из аналогичного другого, что, в общем-то, является рекурсией. Но искусственные узоры, создаваемые человеком, обычно получаются из одинаковых серийно выпускаемых элементов, и в их основе лежит арифметическая прогрессия – каждый элемент подобен другому элементу, но не обязательно всему целому.

12.3. ЛИНЕЙЧАТЫЕ ИЗОБРАЖЕНИЯ

Важным классом узоров являются линейчатые изображения на плоскости.

Эти геометрические конструкции также задаются конечным набором дискретных параметров. Предлагаемая ниже модель позволяет создавать и анализировать регулярные узоры и орнаменты. В ряде предметных областей (кристаллография, распознавание образов, теория упаковок) рассматриваются объекты похожие на узоры.

Основой математической модели построения регулярных узоров является обобщение понятия многоугольников [19].

12.3.1. Обобщенные многоугольники

Рассмотрим многоугольники на плоскости. Как известно, выпуклый многоугольник задается системой линейных неравенств

$$(a_i, x) \leq b_i, a_i \in R^2, i = 0, \dots, n-1. \quad (12.1)$$

Будем предполагать, что система (12.1) не содержит избыточных ограничений. Контур плоского выпуклого многоугольника можно обойти, например, по часовой стрелке. Некоторые алгоритмы, например симплекс-метод, используют именно контур или ребра для обхода вершин. Пронумеруем ребра по обходу контура так, что i -е ребро лежит на i -м ограничении системы (12.1).

Под совокупностью подсистем линейных ограничений понимается объединение множеств решений этих подсистем. Тогда контур многоугольника (12.1) можно задать в виде совокупности S подсистем L_i вида

$$(L_i): \begin{cases} (a_{i-1}, x) \leq b_{i-1}, \\ (a_i, x) = b_i, \\ (a_{i+1}, x) \leq b_{i+1}. \end{cases}$$

Индексация векторов a_i осуществляется по модулю n . Например, подсистема L_0 задается векторами a_{n-1} , a_0 и a_1 . Таким образом, любое ограничение из (12.1) входит в подсистемы L_i трижды: два раза в виде неравенства и один раз в виде уравнения.

Характерной особенностью выпуклого контура является то, что любое неравенство из (12.1) при переходе от одной подсистемы L_i к следующей не меняет знака. Используем это свойство для обобщения понятия выпуклого контура.

Предположим, что знаки неравенств исходной системы (12.1) выбраны произвольно и зафиксированы. Допустим, ограничения подсистем L_i совокупности S сохраняют эти знаки. Тогда такой контур мы будем называть квазивыпуклым. Некоторые ребра контура могут стать невидимыми.

Рассмотрим часть многоугольника, состоящую из трех смежных ребер, задаваемых системами L_{i-1} , L_i , L_{i+1} и четырех последовательных вершин A , B , C , D , лежащих на этих ребрах. Если третье ограничение системы L_{i-1} и первое ограничение системы L_{i+1} имеют разные знаки, то ребро BC будет задавать перегиб контура. Это означает, что вершины A и D будут лежать по разные стороны от прямой, проходящей через ребро BC . Обобщенно выпуклые контуры не имеют ребер перегиба.

Откажемся теперь от требования обобщенной выпуклости, т. е. будем предполагать, что знаки неравенств в ограничениях соседних подсистем L_{i-1} , L_i не зависят друг от друга. Объединение множеств решений всех подсистем L_i будем называть обобщенным многоугольником. Примерами обобщенных многоугольников являются несвязанные ломаные, например текст, буквы которого составлены из отрезков. Если знаки неравенств подсистем L_i заменить на равенства, то этот текст превращается в хаотический набор прямых линий и неузнаваем на этом фоне. Предлагаемый способ позволяет также рисовать картинки и прятать изображение в сетке линий. Ключом для дешифрования изображения являются две последовательности знаков неравенств. Одна последовательность формируется по первым ограничениям подсистем L_i , а вторая по третьим ограничениям.

Для численного моделирования обобщенных многоугольников было разработано приложение на языке C++ под *DOS*, которое, в частности, позволило представить фразу «УГТУ-85» с помощью обобщенного многоугольника (рис. 12.4).

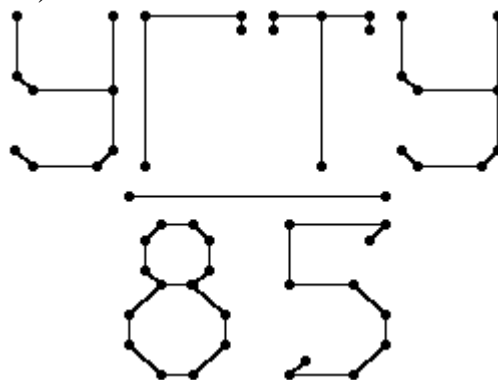


Рис. 12.4. Пример обобщенного многоугольника в виде текста

Если удалить информацию о знаках неравенств, то все видимые и невидимые ребра этого многоугольника будут заменены содержащими их прямыми. В полученном хаосе (рис. 12.5) угадать исходный текст практически невозможно.

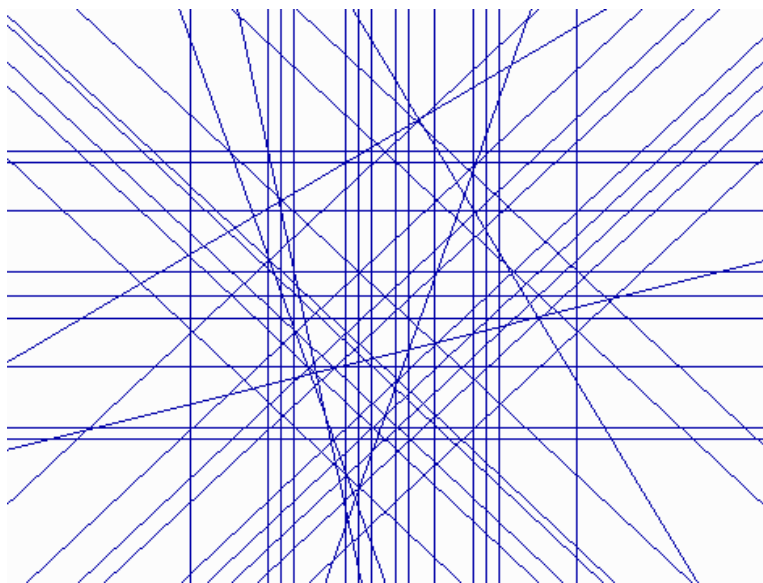


Рис. 12.5. Скрытый текст за хаосом прямых линий

Введение понятия обобщенных многоугольников позволяет экономить память для хранения информации о контуре. Действительно, для каждой подсистемы L_i достаточно двух бит для хранения знаков первого и третьего ограничений. Таким образом, в 1 кБ можно хранить информацию о 4000 обобщенных ребрах. При этом коэффициенты самих линий кодировать не надо. Сетка линий может быть открыта.

12.3.2. Аналитический подход к построению линейчатых узоров

Рассмотрим математическую модель, которая позволяет строить разнообразные плоские узоры, состоящие из отрезков прямых [20].

Возьмём на плоскости прямую L , заданную уравнением $ax + by = c$.

Дополним её тройкой вещественных чисел $\Delta = (\Delta_a, \Delta_b, \Delta_c) \neq 0$, которые будут задавать приращения коэффициентов прямой L .

Определение 1. Счётное множество $F(L, \Delta)$ прямых, коэффициенты которых отличаются относительно коэффициентов базовой прямой L на величину, кратную Δ , назовём семейством прямых

$$F(L, \Delta) = \{(a + i\Delta_a)x + (b + i\Delta_b)y = c + i\Delta_c : i \in \mathbb{Z}\}.$$

Таким образом, прямые из семейства $F(L, \Delta)$ можно индексировать: F_{-1}, F_0, F_1, \dots . Если $\Delta_a = \Delta_b = 0$, то прямые из семейства F параллельны друг другу. Если $\Delta_c = 0$, то прямые имеют общую точку.

Семейство прямых $F(L, \Delta)$ будем называть семейством параллельных прямых, если $\Delta_a = \Delta_b = 0$.

Возьмём три семейства параллельных прямых F , G и H такие, чтобы прямые из F и G , а также из G и H не были параллельны друг другу. В каждом семействе зафиксируем по одной прямой соответственно с целыми индексами i' , j' и k' . Вектор $\vec{z} = (i', j', k')$ назовём вектором начальных индексов. Обозначим через \vec{z} направленный отрезок на прямой $G_{j'}$, ограниченный прямыми $F_{i'}$ и $H_{k'}$.

Пересечение прямых $F_{i'}$ и $H_{k'}$ будем называть левым концом отрезка \vec{z} , а общую точку прямых $G_{j'}$ и $H_{k'}$ – правым концом \vec{z} .

Зафиксируем несколько ненулевых целочисленных векторов $\delta_t = (\delta_{t,1}, \delta_{t,2}, \delta_{t,3})$, $t = 1, \dots, n$. Векторы δ_t будут являться смещениями координат i' , j' и k' вектора \vec{z} . Обозначим через Δ матрицу с размерами $3 \times n$, состоящую из столбцов δ_t , $t = 1, \dots, n$. Из смещений δ_t можно составлять линейные целочисленные комбинации.

$$\Delta \cdot \alpha = \sum_{i=1}^n \alpha_i \cdot \Delta_i, \text{ где } \alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n.$$

Таким образом, векторы δ_t , $t = \overline{1, n}$, порождают линейное подпространство в \mathbb{Z}^3 над кольцом целых чисел.

Рассмотрим направленный отрезок $\overrightarrow{z + \delta}$, $\delta = \Delta \cdot \alpha$, который лежит на прямой $G_{j' + \delta_2}$, ограниченный прямыми $F_{i' + \delta_1}$ и $H_{k' + \delta_3}$. Если $F = H$ и $\delta_{t,1} = \delta_{t,3}$ для всех t , то отрезки $\overrightarrow{z + \delta}$, кроме того, имеют одинаковую длину.

Определение 2. Пусть F , G и H – семейство параллельных прямых таких, что F и H не параллельны G , \vec{z} – вектор начальных индексов, Δ – матрица с размерами $3 \times n$, составленная из вектор-столбцов целочисленных смещений. Элементарным линейчатым узором $U(F, G, H, \vec{z}, \Delta)$ называется бесконечное множество параллельных отрезков возможно различной длины вида $\overrightarrow{z + \Delta\alpha}$, $\alpha \in \mathbb{Z}^n$, полученных ограничением прямых из семейства G прямыми из семейств F и H . Отрезок \vec{z} назовём ядром узора U .

Таким образом, элементарный линейчатый узор задаётся тройкой

параллельных семейств, тройкой индексов ядра узора и несколькими тройками смещений индексов ядра узора. Общее количество данных составляет: $3 \times 3 = 9$ вещественных коэффициентов базовых прямых семейств, $3 \times 3 = 9$ вещественных коэффициентов приращений семейств (среди которых шесть коэффициентов равны 0); 3 целых индекса ядра и $3 \times n$ целых смещений индексов. В итоге для задания элементарного линейчатого узора достаточно 12 вещественных чисел и $3 \cdot (n + 1)$ целых чисел, где $n \geq 1$ – целое.

Совмещая несколько элементарных линейчатых узоров, мы получим более сложные и информативные изображения.

Определение 3. Объединение нескольких элементарных линейчатых узоров назовём регулярным линейчатым узором.

Таким образом, регулярные узоры лежат на 3 или более семействах параллельных прямых.

ПРИМЕР 12.2. Возьмем три семейства прямых (рис. 12.6):

- семейство А, заданное прямой $x + y = 0$ и смещением коэффициентов прямой $\Delta_A = \{0, 0, 10\}$;
- семейство В, заданное прямой $-x + y = 0$ и смещением коэффициентов прямой $\Delta_B = \{0, 0, 10\}$;
- семейство С, заданное прямой $y = 0$ и смещением коэффициентов прямой $\Delta_C = \{0, 0, 5\}$.

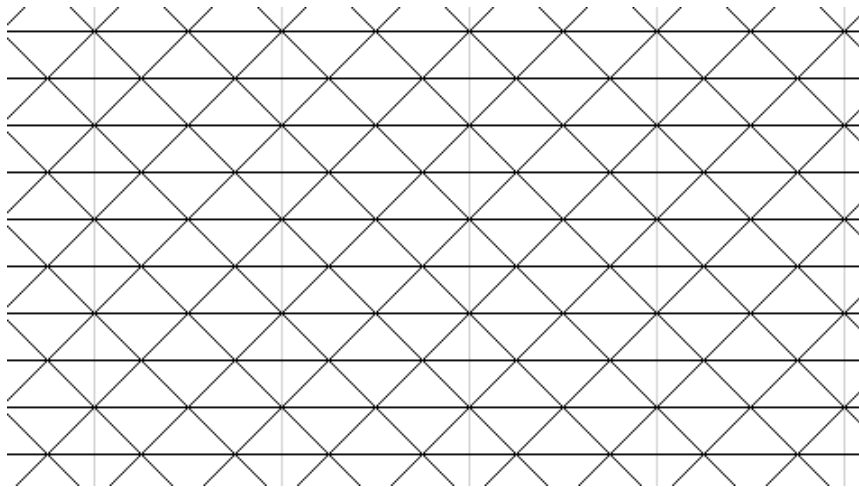


Рис. 12.6. Три семейства прямых регулярного узора

Построим элементарный узор U_1 на семействах $F = A$, $G = B$, $H = A$. В качестве ядра z элементарного узора U_1 выберем отрезок, лежащий на прямой B_0 и ограниченный прямыми A_0 и A_1 . Ядро задаётся прямыми A_0 , B_0 , A_1 , т. е. определяется начальными индексами $z = (0, 0, 1)$. Зададим для ядра смещения индексов $(6, 0, 6)$ и $(2, 2, 2)$, т. е.

$$\Delta = \begin{pmatrix} 6 & 2 \\ 0 & 2 \\ 6 & 2 \end{pmatrix}.$$

Результатом будет элементарный узор на рис. 12.7.



Рис. 12.7. Элементарный узор U_1

Построим еще 7 элементарных узоров.

$U_2: F = A, G = B, H = A, z = (2, 0, 5)$, смещения $(6, 0, 6), (2, 2, 2)$;

$U_3: F = A, G = B, H = A, z = (0, 1, 2)$, смещения $(3, 0, 3), (2, 2, 2)$;

$U_4: F = B, G = A, H = B, z = (1, 0, -1)$, смещения $(3, 0, 3), (2, 2, 2)$;

$U_5: F = B, G = A, H = B, z = (0, 1, 1)$, смещения $(6, 0, 6), (2, 2, 2)$;

$U_6: F = B, G = A, H = B, z = (0, 1, 1)$, смещения $(2, 0, 5), (2, 2, 2)$;

$U_7: F = B, G = C, H = B, z = (0, 1, -1)$, смещения $(6, 0, 6), (-2, 2, -2)$;

$U_8: F = B, G = C, H = B, z = (-2, 1, -5)$, смещения $(6, 0, 6), (-2, 2, -2)$;

$U_9: F = B, G = C, H = B, z = (2, 2, 0)$, смещения $(3, 0, 3), (1, 2, 1)$.

В результате объединения 9 элементарных узоров получим регулярный узор (рис. 12.8). Подобным образом происходит укладка тротуарных плиток.

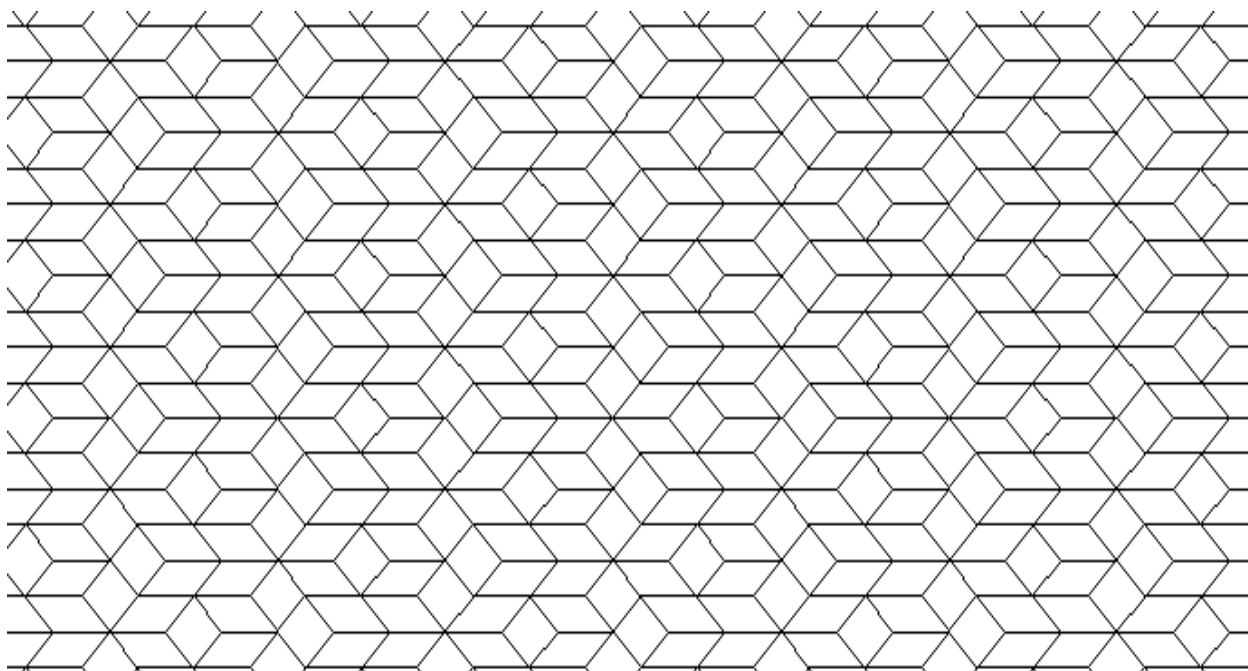


Рис. 12.8. Фрагмент регулярного линейчатого узора

На рис. 12.9 показана схема наложения элементарных узоров. Можно заметить, как 9 узоров сходятся в одну область, в которой образуется регулярный узор.

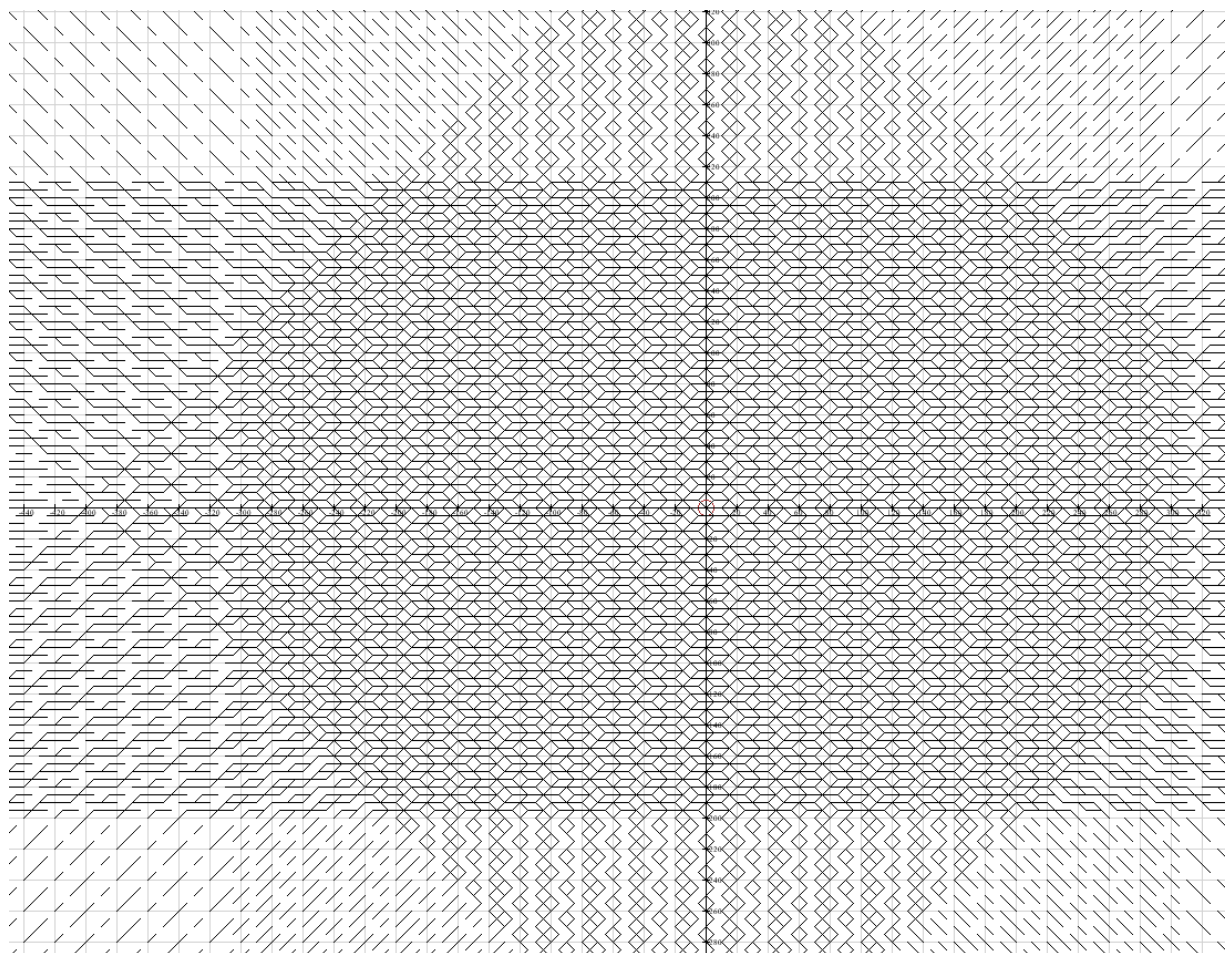


Рис. 12.9. Наложение 9 элементарных узоров

Для распознавания ряда узоров, в частности «тротуарной плитки» и «пчелиных сот» с наложенным шумом, используется программный комплекс [21]. Для анализа узоров использовались алгоритмы из [22]. Результаты работы могут быть использованы для сжатия изображения, шифрования данных или определения степени регулярности изображения.

В заключение отметим, что аналогичные исследования узоров и орнаментов проводятся в области архитектуры [23], в легкой промышленности, в частности при проектировании ковровых узоров [24].

Плоские регулярные узоры в форме шестиугольных «пчелиных сот» являются геометрической моделью одноатомной углеродной кристаллической решетки, получившей название графен [25]. Перспективы этого материала бурно обсуждаются в последнее время, и за открытие и исследование графена Андрей Гейм и Константин Новосёлов получили Нобелевскую премию за 2010 год.

Регулярные узоры тесно связаны с плоскими и трехмерными точечными узорами [26]. С помощью новых методов и алгоритмов [27] можно построить и исследовать периодические нормальные упаковки невыпуклых многогранников (поликубов) в пространствах произвольной размерности.

13. ОЛИМПИАДНЫЕ ЗАДАЧИ И ИХ РЕШЕНИЯ⁷

Задача «Диверсантки»

Некое государство обратилось к России с просьбой о помощи в борьбе с терроризмом. С этой целью правительство предложило товарищу Ф.И. Сухову подобрать команду женщин-диверсанток для заброски в тыл врага. Сухов послал Петруху для сбора информации о жительницах вражеской столицы. Петруха, переодевшись уличным торговцем, тщательно записывал вес и рост женщин. Помогите Сухову собрать команду под руководством Гюльчатай, которая не вызовет подозрений у террористов по весу и росту диверсанток.

Решение задачи «Диверсантки»

Информацию, которую собрал Петруха, используем для задания выборочной функции распределения $F(x, y)$ двумерной случайной величины (рост, вес) следующим образом:

$$F(x, y) = P(\text{рост} < x, \text{вес} < y).$$

Функция F определена в области $[a, b] \times [c, d]$. Определим по функции F функцию распределения G первой составляющей «рост» $G(x) = F(x, d)$.

В соответствии с алгоритмом «обратной функции» генерируем случайную величину x_1 , распределенную по закону G :

$$x_1 = G^{-1}(z),$$

где z – равномерно распределенная случайная величина на отрезке $[a, b]$. По выбранному x_1 строим условное распределение $H(y)$ второй составляющей «вес»

$$H(y) = F(x_1, y) / F(x_1, d).$$

В соответствии с алгоритмом «обратной функции» генерируем случайную величину y_1 , распределенную по закону H :

$$y_1 = H^{-1}(w),$$

где w – равномерно распределенная случайная величина на отрезке $[c, d]$.

В результате получили новую реализацию (x_1, y_1) , по которой подбираем первую кандидатуру в команду. Алгоритм повторяем несколько раз.

Задача «Революция»

Жили-были два государства A и B . Каждый гражданин с рождения получал раз и навсегда фиксированный размер ежемесячного дохода, случайно равномерно распределенный на интервале от 1 до 9 у.е. И было у них богатых столько же, сколько бедных. Богатыми считались люди, доход которых входит в верхний дециль, то есть в верхние 10 % диапазона доходов. У бедных – доход из нижнего дециля.

В обеих странах произошли революции. В стране A решили выдавать вместо прежней суммы дохода корень квадратный из этого дохода и назвали новую денежную единицу скрутлик ($\sqrt{}$). В стране B стали выдавать квадрат

⁷ В данной главе представлены авторские (С.П. Трофимов) задачи и их решения для олимпиады по специальности «Вычислительные машины, комплексы, системы и сети» [28], проходящей с 2000 года на кафедре автоматики и информационных технологий «Уральского федерального университета имени первого Президента России Б.Н. Ельцина».

прежней суммы соответственно в скурликах (*sqr*). Валюты перестали конвертироваться.

Определите в странах *A* и *B* долю богатых и бедных. Какую революцию можно считать коммунистической, а какую буржуазной?

Решение задачи «Революция»

Пусть X – случайная величина с плотностью распределения $f(x)$, $\varphi(x)$ – монотонно возрастающая функция. Рассмотрим новую случайную величину $Y = \varphi(X)$. Обозначим через $\psi(y)$ функцию, обратную к $y = \varphi(x)$. Тогда функция распределения величины Y имеет вид

$$G(Y) = P(Y < y) = P(X < \psi(y)) = \int_{-\infty}^{\psi(y)} f(x) dx.$$

Дифференцируя данное выражение, получаем плотность распределения Y

$$g(y) = G'(y) = f(\psi(y)) \psi'(y).$$

В нашем случае $f(x) = 1/8$ для $x \in [1, 9]$.

Для страны *A* имеем $\varphi(x) = \sqrt{x}$, $\psi(y) = y^2$, $\psi'(y) = 2y$.

$g_A(y) = y/4$ для $y \in [1, 3]$.

Для страны *B* имеем $\varphi(x) = x^2$, $\psi(y) = \sqrt{y}$, $\psi'(y) = \frac{1}{2\sqrt{y}}$.

$$g_B(y) = \frac{1}{16\sqrt{y}} \text{ для } y \in [1, 81].$$

Обозначим через R разницу между максимальным и минимальным доходами. Бедными считаются те, у кого доход не превышает минимальный более чем на $R/10$. Соответственно богатыми считаются те, у кого доход превышает минимальный более чем на $9/10R$. Подсчитаем их количество.

$$\text{Бедные}(A) = \int_1^{1.2} \frac{y}{4} dy = 5,5 \, \%.$$

$$\text{Богатые}(A) = \int_{2,8}^3 \frac{y}{4} dy = 14,5 \, \%.$$

$$\text{Бедные}(B) = \int_1^9 \frac{1}{16\sqrt{y}} dy = 25 \, \%.$$

$$\text{Богатые}(B) = \int_{73}^{81} \frac{1}{16\sqrt{y}} dy = 5,7 \, \%.$$

Учитывая, что до революций количество бедных и богатых было одинаково, логично считать, что в стране *A* произошла коммунистическая революция, а в стране *B* – буржуазная.

Задача «Зарплата»

Программист устраивается на работу с ежемесячной зарплатой в конце месяца 10 000 руб. Хозяин честно его предупредил, что зарплата выплачивается нерегулярно с задержкой, причем дата выдачи задержанной зарплаты

распределена равномерно в течение двух следующих месяцев. Какую добавку к зарплате следует попросить, чтобы компенсировать эту задержку?

Предположить, что инфляции нет, цена капитала 10 % годовых, срок контракта бесконечен и зарплата будет постоянной. Использовать формулу дисконтирования будущей суммы FV на время t по сложной процентной ставке r

$$PV = FV(1 + r)^{-t},$$

где PV – приведенная сумма на момент контракта.

Решение задачи «Зарплата»

В силу условия задачи все платежи одинаковы, поэтому достаточно рассчитать приведенную сумму одной задержанной зарплаты на дату $t = 0$ ее фактической выплаты.

Обозначим:

m – срок максимальной задержки в месяцах;

r – годовая ставка;

A – исходная зарплата;

B – скорректированная зарплата.

Плотность распределения даты фактической выплаты

$$f(t) = \begin{cases} 12/m & \text{при } t \in [0, m/12], \\ 0 & \text{иначе.} \end{cases}$$

Тогда приведенная сумма PV случайного платежа B на дату 0 равна

$$PV = \int_0^{\infty} Bf(t)(1+r)^{-t} dt = \int_0^{m/12} B(12/m)(1+r)^{-t} dt.$$

После корректировки зарплаты финансовые интересы сторон не должны измениться, поэтому имеем

$$PV = A,$$

откуда

$$B = A \ln(1+r) / [12/m \cdot (1 - (1+r)^{-m/12})] = 10\,079,64 \text{ руб.}$$

Ответ. Добавка к зарплате 79,64 руб.

Задача «Ковер»

Гарем шаха решил сделать подарки к 85-летию своего владыки. Все жены и наложницы были разбиты на команды по семь человек и каждой команде дали одно и то же задание: соткать круглый ковер, состоящий из семи секторов. Каждый сектор ткется одной женщиной и только одним узором из трех возможных. Женщины друг друга не видят и выбирают узор случайно с вероятностью $1/3$.

Ковер считается удачным, если соседние секторы имеют разный узор. Найти вероятность удачного ковра, за который шах щедро наградит команду.

Решение задачи «Ковер»

Пусть:

n – количество женщин в одной команде;

m – количество узоров.

Обозначим $f(m, n)$ количество удачных ковров с m узорами из n секторов.

Пронумеруем секторы и ткачих от 1 до n . Будем рассматривать размещения узоров по секторам. Узоры пронумеруем от 1 до m .

Рассмотрим частный случай

$$g(n) = f(3, n).$$

Найдем рекуррентную формулу для $g(n)$.

При $n = 2$ удачными коврами являются последовательности 12, 13, 21, 23, 31, 32, то есть $g(2) = 6$.

При $n = 3$ удачными коврами являются последовательности 123, 132, 213, 231, 312, 321, т. е. $g(3) = 6$.

Допустим ковер $u_1 u_2 \dots u_{n-2} u_{n-1} u_n$ удачный, где $n \geq 4$ и все $u_i \in \{1, 2, 3\}$. Возможны два взаимоисключающих случая.

1. $u_1 \neq u_{n-1}$. Тогда ковер $u_1 u_2 \dots u_{n-2} u_{n-1}$ будет удачным и подбор узора u_n однозначный – отличный от u_1 и u_{n-1} .

2. $u_1 = u_{n-1}$. Тогда ковер $u_1 u_2 \dots u_{n-2} u_{n-1}$ не будет удачным. Зато ковер $u_1 u_2 \dots u_{n-2}$ будет удачным, так как $u_{n-2} \neq u_{n-1} = u_1$. Тогда узор u_n можно выбрать двумя способами – из трех чисел 1, 2, 3 можно выбрать два числа, отличные от u_{n-1} .

Итак, из одного удачного узора из $(n - 1)$ секторов можно построить один удачный узор из n секторов и из одного удачного узора из $(n - 2)$ секторов можно построить два удачных узора из n секторов. Отсюда имеем

$$g(n) = g(n - 1) + 2g(n - 2). \quad (13.1)$$

Заметим, что формула (13.1) очень похожа на известную рекуррентную формулу для чисел Фибоначчи $F(n)$

$$F(n) = F(n - 1) + F(n - 2).$$

Последовательно вычисляем $g(4) = 18$, $g(5) = 30$, $g(6) = 66$, $g(7) = 126$. Общее количество ковров из 7 секторов, очевидно, равно 3^7 . Тогда искомая вероятность удачного ковра равна $126/3^7 = 5,8 \%$.

Формула (13.1) допускает обобщение

$$f(m, n) = (m - 2) \cdot f(m, n - 1) + (m - 1) \cdot f(m, n - 2). \quad (13.2)$$

Задача «Весы»

Имеются напольные электронные весы с ценой деления 100 г.

Мама решила взвесить маленького ребенка следующим образом. Сначала она взвесилась вместе с ним, затем взвесилась сама и в результате вычитания получила вес ребенка.

Но каково было ее удивление, когда после аналогичной операции с папой у ребенка получился другой вес!

Проясните маме эту ситуацию. Найдите ряд распределения разности двух взвешиваний.

Погрешность дискретизации измерений считать равномерной в интервале от -50 до 50 г.

Решение задачи «Весы»

Пусть M – фактический вес мамы без ребенка, MR – вес мамы с ребенком. По условию задачи M и MR – равномерно распределенные случайные величины на интервалах $[VM - 0,05; VM + 0,05]$ и $[VMR - 0,05; VMR + 0,05]$, где VM и

VMR – вес в килограммах, кратный 50 граммам и показанный на весах. Тогда первый вес ребенка $R1 = MR - M$ является случайной величиной с треугольным распределением на интервале $(VMR - VM) + [-0,10; 0,10]$.

Пусть P – фактический вес папы без ребенка, PR – вес папы с ребенком. По условию задачи P и PR также равномерны на интервалах $[VP - 0,05; VP + 0,05]$ и $[VPR - 0,05; VPR + 0,05]$, где VP и VPR – вес в килограммах, кратный 50 граммам. Тогда второй вес ребенка $R2 = PR - P$ – случайная величина с таким же треугольным распределением на интервале $(VPR - VP) + [-0,10; 0,10]$.

Предположим, что $VMR - VM = VPR - VP$. Тогда случайные величины $R1$ и $R2$ имеют одинаковое распределение и их разность $dR = R1 - R2$ является разностью двух треугольных распределений. Множеством реализаций величины dR является отрезок $[-0,20; 0,20]$. Обозначим плотность dR через $fdR(x)$. Эта функция является сверткой плотностей величин $R1$ и $R2$. Округлим dR до 50 граммов. Получим дискретную величину d , для которой надо найти закон распределения. Подсчитаем вероятности:

$P(d = -0,20)$ = интеграл от $fdR(x)$ на интервале $[-0,20; -0,175]$;

$P(d = -0,15)$ = интеграл от $fdR(x)$ на интервале $[-0,175; -0,125]$;

$P(d = -0,10)$ = интеграл от $fdR(x)$ на интервале $[-0,125; -0,075]$;

$P(d = -0,05)$ = интеграл от $fdR(x)$ на интервале $[-0,075; -0,025]$;

$P(d = 0)$ = интеграл от $fdR(x)$ на интервале $[-0,125; 0,125]$.

Аналогично находятся вероятности для положительных значений величины d . Аналитические выражения для данных вероятностей довольно громоздки, поэтому путем моделирования программным способом найдем выборочный ряд распределения

d	-0,20	-0,15	-0,10	-0,05	0	0,05	0,10	0,15	0,20
Вероятность, %	0	1,35	9,90	22,85	31,80	22,85	9,90	1,35	0

Задача «Забывчивый студент»

Забывчивый студент готовится к экзамену по списку вопросов. Вопросы записаны на карточках, которые лежат в некоем автомате. На одной карточке – один вопрос. Сколько карточек – студент не знает, какие вопросы уже повторил – не помнит. Автомат выдает по одной карточке. В каждой руке студент может держать не более одной карточки. Отложенная карточка убирается в недоступное место.

Предложите алгоритм, по которому будет выбран вопрос для повторения. Распределение вопросов для повторения должно быть равномерным, т.е. все вопросы должны повторяться с одинаковой вероятностью. После повторения выбранного вопроса все карточки возвращаются обратно в автомат.

Решение задачи «Забывчивый студент»

Выберем правой рукой первую, взятую наугад карточку. Переложим ее в левую руку. Будем считать, что карточка, которую необходимо повторить, находится в левой руке. Возьмем правой рукой вторую карточку. Случайным равномерно распределенным образом выберем число из множества $\{1, 2\}$. Если

выпала 1, то вместо второй карточки в правую руку берем третью. Если выпала 2, то заменяем карточку, которую необходимо повторить, т.е. перекладываем карточку из правой руки в левую.

Допустим, студент перебрал $(N - 1)$ карточек и у него в левой руке есть карточка с вопросом для подготовки. Студент берет N -ю карточку и выбирает случайное, равномерно распределенное число из множества $\{1, \dots, N\}$. Если выпало N , то меняем карточку в левой руке на N -ю карточку. В противном случае продолжаем выполнять алгоритм до окончания набора карточек.

Доказательство алгоритма

Допустим, что всего имеется N вопросов-карточек. Вероятность $P(N)$ выбора последнего вопроса, очевидно, равна $1/N$. Вероятность $P(N - 1)$ выбора $(N - 1)$ вопроса определим по формуле условной вероятности.

$P(N - 1) = P(N - 1 \text{ при условии, что вопрос из множества } \{1, \dots, N - 1\}) \cdot P(\{1, \dots, N - 1\}) = 1/(N - 1) \cdot (N - 1)/N = 1/N$.

Аналогично имеем

$P(N - 2) = P(N - 2 \text{ при условии, что вопрос из множества } \{1, \dots, N - 2\}) \cdot P(\{1, \dots, N - 2\}) = 1/(N - 2) \cdot (N - 2)/(N - 1) \cdot (N - 1)/N = 1/N$.

И так далее. Наконец,

$P(1) = 1/2 \cdot 2/3 \cdot (N - 1)/N = 1/N$.

Задача «Оценка»

Студент отвечает на экзамене на два вопроса. Ответы на каждый вопрос преподаватель оценивает по непрерывной шкале от 2 до 5. Затем эти две промежуточные оценки усредняются и округляются до ближайшего целого числа.

Свои знания по первому вопросу студент оценивает так. Вероятность получить у преподавателя оценку от 2 до 3 такая же, как вероятность получить оценку от 3 до 5, то есть по 50 %.

Студент уверен, что по второму вопросу получит промежуточную оценку от 3 до 4.

Считая, что все распределения в указанных диапазонах равномерные, найти вероятность получения итоговой оценки: 2, 3, 4 и 5.

Решение задачи «Оценка»

Пусть X и Y – случайные промежуточные оценки за 1 и 2 вопросы, $f(x)$, $g(y)$ их плотности распределения. Тогда $f(x) = 0,5$ на отрезке $[2, 3]$ и $f(x) = 0,25$ на отрезке $[3, 5]$. Также $g(y) = 1$ на отрезке $[3, 4]$.

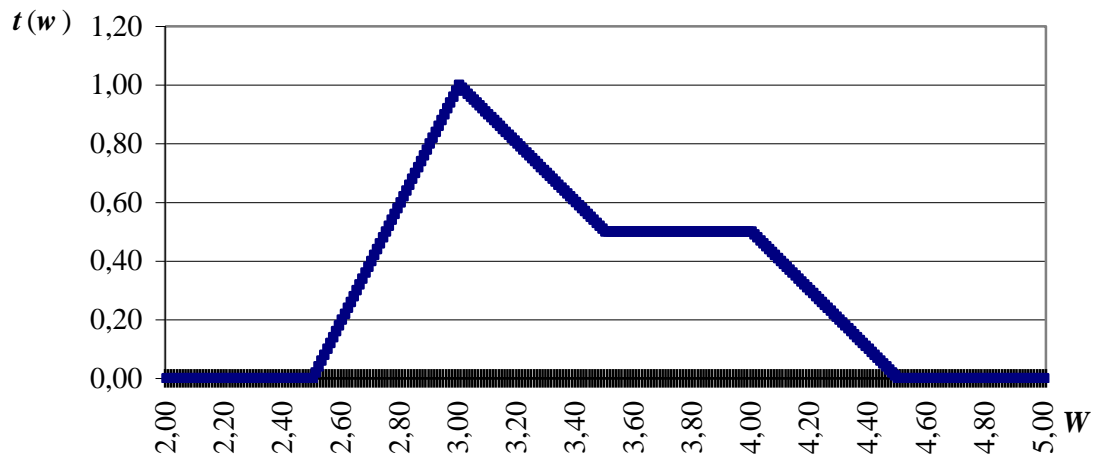
Суммарная оценка $Z = X + Y$ имеет плотность $h(z)$. Плотность $h(z)$ является сверткой функций $f(x)$ и $g(y)$, то есть $h(z)$ является определенным интегралом функции $f(x) \cdot g(z - x)$ по dx по всем x .

Последовательно получаем.

1. При $z < 5$ $h(z) = 0$.
2. При $5 < z < 6$ $h(z) = (z - 5)/2$.
3. При $6 < z < 7$ $h(z) = 2 - z/4$.
4. При $7 < z < 8$ $h(z) = 1/4$.
5. При $8 < z < 9$ $h(z) = (9 - z)/4$.

6. При $9 < z$ $h(z) = 0$.

Осредненная оценка $W = Z/2$ изменяется в интервале от 2,5 до 4,5. Плотность распределения $t(w)$ (см. рисунок) величины W равна $t(w) = 2 \cdot h(z/2)$.



Плотность распределения $t(w)$ величины W

Отсюда видно, что округленная оценка не может равняться 2 или 5.

Для определения вероятности получения «3» подсчитаем интеграл $t(w)$ на интервале от 2,5 до 3, 5. Он равен $5/8$.

Для определения вероятности получения «4» подсчитаем интеграл $t(w)$ на интервале от 3,5 до 4, 5. Он равен $3/8$.

Ответ. Искомые вероятности равны 0; $5/8$; $3/8$; 0.

Задача «Рыбалка-2009»

Два рыбака Иван и Петр владеют каждый своим участком реки. У обоих есть информация о распределении вероятностей поймать рыбу на своем участке. Иван использует свою информацию и каждый раз накануне рыбалки генерирует выбор места рыбалки в соответствии со своим законом распределения.

Петр же ленится это делать.

Предложите алгоритм, по которому для каждой реализации места рыбалки у Ивана можно получить реализацию места для Петра.

Решение задачи «Рыбалка-2009»

Пусть X и Y – случайные величины, характеризующие наличие рыбы на участках Ивана и Петра. Обозначим $F(x)$ и $G(y)$ их функции распределения.

Известно, что случайную величину X можно получить из СВ Z , равномерно распределенной на отрезке $[0, 1]$, методом «обратной функции»: $X=F^{-1}(Z)$. В данном случае наоборот, СВ Z мы строим с помощью X , а затем из Z получаем Y .

Итак, $Z=F(X)$.

В результате получаем следующий алгоритм:

$Y = G^{-1}(Z) = G^{-1}(F(X))$.

Задача «Любит – не любит - 2010»

Имеется две одномерные, непрерывные случайные величины X и Y с известными законами распределения. Известно, что над ними можно выполнять различные операции: сложение, умножение, возведение в степень. Методом «обратной функции» можно генерировать реализации случайной величины.

Предлагается новая бинарная операция $\#$. Реализации случайной величины $Z = X \# Y$ строятся по правилу чередования: четные номера берутся в соответствии с X , нечетные – в соответствии с Y .

Найдите закон распределения Z .

Решение Задача «Любит – не любит - 2010»

Пусть $f(x)$, $g(y)$ и $h(z)$ – функции плотностей распределения случайных величин X , Y и Z соответственно. Функции f и g известны, найдем h .

Оценим вероятность попадания Z в маленький интервал dz . По формуле полной вероятности с учетом чередования получаем

$$\begin{aligned} P(z \leq Z \leq z + dz) &= h(z) dz = 1/2 \cdot P(z \leq X \leq z + dz) + 1/2 \cdot P(z \leq Y \leq z + dz) = \\ &= 1/2 \cdot f(z) dz + 1/2 \cdot g(z) dz. \end{aligned}$$

Отсюда после сокращения на произвольное приращение dz получаем ответ $h(z) = 1/2 \cdot f(z) + 1/2 \cdot g(z)$.

Задача обобщается следующим образом. Пусть имеется однопараметрическое семейство случайных величин $V(\alpha)$ с плотностями распределения $f(x, \alpha)$. Выберем параметр α_0 в соответствии с известной плотностью $g(\alpha)$ и в качестве реализации неизвестной случайной величины Z возьмем реализацию известной случайной величины $V(\alpha_0)$.

Тогда функция плотности распределения случайных величин Z

$$h(z) = \int_{-\infty}^{+\infty} g(\alpha) f(z, \alpha) d\alpha.$$

14. ИНФОРМАЦИОННО-КОММУНИКАЦИОННАЯ ТРАНСПОРТНАЯ СИСТЕМА

14.1. ЛОГИСТИЧЕСКАЯ МОДЕЛЬ ТРАНСПОРТНОГО ПРЕДПРИЯТИЯ НА ОСНОВЕ ИНФОРМАЦИОННО-КОММУНИКАЦИОННОЙ СИСТЕМЫ

Транспортный процесс перевозки пассажиров можно представить в виде технической системы. Входом системы является потребность населения в перевозках и наличие определенного числа подвижного состава. Выходом системы является своевременная доставка пассажиров в пункты назначения. Обратная связь заключается в поступлении с линии информации о движении подвижного состава, соблюдении расписания, интервалов движения и соответствии числа подвижного состава потребностям в перевозках. Ограничения – это соблюдение заданного скоростного режима, обеспечение безопасности и комфортности поездок и выполнение финансовых показателей. Процесс функционирования системы решает проблему безопасной и своевременной доставки пассажиров в пункты назначения. Идеальная система

имеет нулевую разницу между существующей и заданной системами. Навигационная система, как наблюдатель, в реальном времени отслеживает соблюдение расписания, позволяет оперативно управлять транспортным процессом. В качестве критериев оценки эффективности транспортного потока принимают безопасность движения, оптимальное использование вместимости подвижного состава и снижение затрат на перевозку.

Анализ перевозочного процесса и эффективности работы транспортного предприятия с использованием логистической модели информационно-коммуникационной системы позволяет качественно управлять процессом перевозки пассажиров на основе отчетных технико-экономических показателей о работе подвижной единицы, водителя и кондуктора, структурных подразделений и всего предприятия в целом.

Круг проблем, которые затрагивает логистика, – это вопросы управления различными потоками: материальными, информационными, экономическими. Основная задача логистики – предоставить потребителю услуги (товар) в заданное время заданного качества в заданное место и за определенную цену.

Логистика отвечает за две составляющие стоимости услуги (товара) – время и место. Логистика обеспечивает наличие услуги (продукта) тогда и там, когда и где это нужно потребителям, а с помощью маркетинга определяются свойства услуги (продукта), знакомят с ней потребителей (в виде расписания и навигационной системы через Интернет, сотовую связь, информационные табло на остановочных комплексах), и разрабатывается механизм заключения сделок по оплате услуг (абонемент, электронная система оплаты проезда – Е-карта – персональная (социальная) и не персональная). Основная функция пассажирской транспортной логистики – это управление транспортными потоками по графику в зависимости от потребностей пассажиров. Транспорт представляет собой важное звено логистической системы; он должен обладать рядом необходимых свойств и удовлетворять определенным требованиям в целях создания инновационных систем сбора и распределения пассажиров. Главное свойство транспорта – способность перевозить небольшие потоки пассажиров через короткие интервалы времени в соответствии с меняющимися потребностями пользователя.

Показатели логистики – безопасность движения, время перевозки пассажиров, точность выполнения расписания (соблюдение графика движения), бесперебойность, готовность к оказанию услуг по перевозке, качество услуг по перевозке (определяется количеством пассажиров, перевезенных по расписанию), гибкость (готовность предприятия выполнить новые услуги по требованию государственных органов и пассажиров), доступность информации (способность предприятия выдавать запрашиваемые клиентом сведения о расписании и движении пассажирского транспорта), согласование (стыковка расписания движения различных маршрутов и разных видов транспорта).

Сущность логистики заключается в комплексе предоставляемых услуг – удобное расписание, комфортный транспорт, регулярные перевозки, приемлемая цена за перевозку. Применяя логистические методы можно добиться существенного снижения уровня затрат.

На рис. 14.1 представлена логистическая модель транспортного предприятия, состоящая из основных потоков управления: материального (перевозка пассажиров), информационного, экономического (доход депо и предприятия в целом).



Рис. 14.1. Логистическая модель транспортного предприятия

Условием работы логистической модели является потребность пассажиров в перевозке и наличие самого транспортного предприятия. Определяющими факторами формирования маршрутной сети являются направления, распределение по территории и мощность пассажирских потоков. Отчетно-статистический метод обследования мощности пассажирских потоков опирается на количество проданных абонементов и транзакций (количества

поездов), совершенных по электронным Е-картам. Автоматизация информационных потоков на основе оригинальной информационно-коммуникационной системы службы автоматики и связи МУП ТТУ г. Екатеринбурга [29–35] обеспечивает своевременное получение статистической информации.

Информационно-коммуникационная система МУП ТТУ г. Екатеринбурга (рис. 14.2) включает в себя следующие взаимосвязанные программные комплексы: «Составление расписания маршрутизированного транспорта», «Подготовка нарядов водителей и кондукторов», «Диспетчер выпуска и движение подвижной единицы (ПЕ)», «Табель учета рабочего времени водителей и кондукторов», «Путевой лист автотранспортной службы МУП ЕТТУ», «Транспорт города Екатеринбурга» [29–35].



Рис. 14.2. Модель информационно-коммуникационной системы ЕМУП ТТУ

Данная система позволяет получить отчетные технико-экономические показатели о работе ПЕ, водителя и кондуктора, структурных подразделений и

всего предприятия в целом. Для реализации информационно-коммуникационной системы на Web-сервере разработана единая база данных, позволяющая оперативно и полно интегрировать данные из одного модуля в другой, не дублируя информацию.

Информационно-коммуникационная система реализована в современной архитектуре клиент-сервер, серверная часть функционирует в операционной среде Windows NT или LINUX с использованием СУБД MySQL, клиентские места работают под управлением Windows 9x, Windows NT и выше. Для работы с базой данных на сервере должна быть установлена СУБД MySQL 5.0.45, а на клиентской машине – BDE Administrator Delphi 2007.

С помощью программного комплекса «Составление расписания маршрутизированного транспорта» составляется расписание движения трамваев и троллейбусов, с возможностью формирования на любые дни недели, на конечные станции, промежуточные контрольные пункты, остановочные пункты и для каждого наряда (поездные).

Программный комплекс «Составление расписания маршрутизированного транспорта» предусматривает:

- автоматический ввод данных для расчета;
- работу с электронной маршрутной схемой;
- составление и редактирование расписания;
- автоматизированную корректировку поездных расписаний одновременно с маршрутным расписанием;
- увязку расписаний по участкам маршрутной сети;
- формирование наряда выпуска подвижного состава в целом по сети, в т.ч. и развернутого наряда по маршрутам, графикам и сменам.

После составления расписания маршрутизированного электротранспорта производится расчет технико-экономических показателей планируемой работы подвижного состава, маршрутов, водителей и предприятия в целом, в том числе:

- количество рейсов по каждому наряду, парку, маршруту и в целом по сети;
- наличие подвижного состава по часам суток по каждому маршруту и в целом по сети;
- интервалы движения, вагоно-часы, вагоно-километры.

Правильно составленные с помощью данной программы расписания обеспечивают выполнение основных плановых показателей по регулярности движения пассажирского транспорта, перевозке пассажиров и выручке.

Программный комплекс «Составление расписания маршрутизированного транспорта» выполняет следующие функции:

- хранение набора расписаний;
- создание нового расписания (в том числе на основе одного из предыдущих);
- просмотр и модификация исходных данных для составления расписания;
- автоматическое формирование поездного расписания;
- автоматический расчет плановых показателей;
- составление групп для наряда;

- формирование справок и дополнительных отчетов;
- подготовка итоговых выходных данных для использования другими программами.

Расчет технико-экономических показателей обеспечивает возможность формирования действующих и создания новых выходных форм в соответствии с требованиями предприятия, корректировки расписаний движения маршрутов по конечным станциям, контрольным пунктам, остановкам и поездных расписаний. Реализована возможность просмотра и печати всех формируемых документов.

На рис. 14.3 представлена концептуальная модель «Составление расписания маршрутизированного транспорта». Для каждого вида электротранспорта формируется своя схема движения – электронная визуализированная маршрутная сеть с автоматизированной корректировкой расписания (рис. 14.4). Кроме того, выбрав маршрут, можно увидеть его на общей схеме транспортной сети (рис. 14.5).



Рис. 14.3. Концептуальная модель «Составление расписания маршрутизированного транспорта»

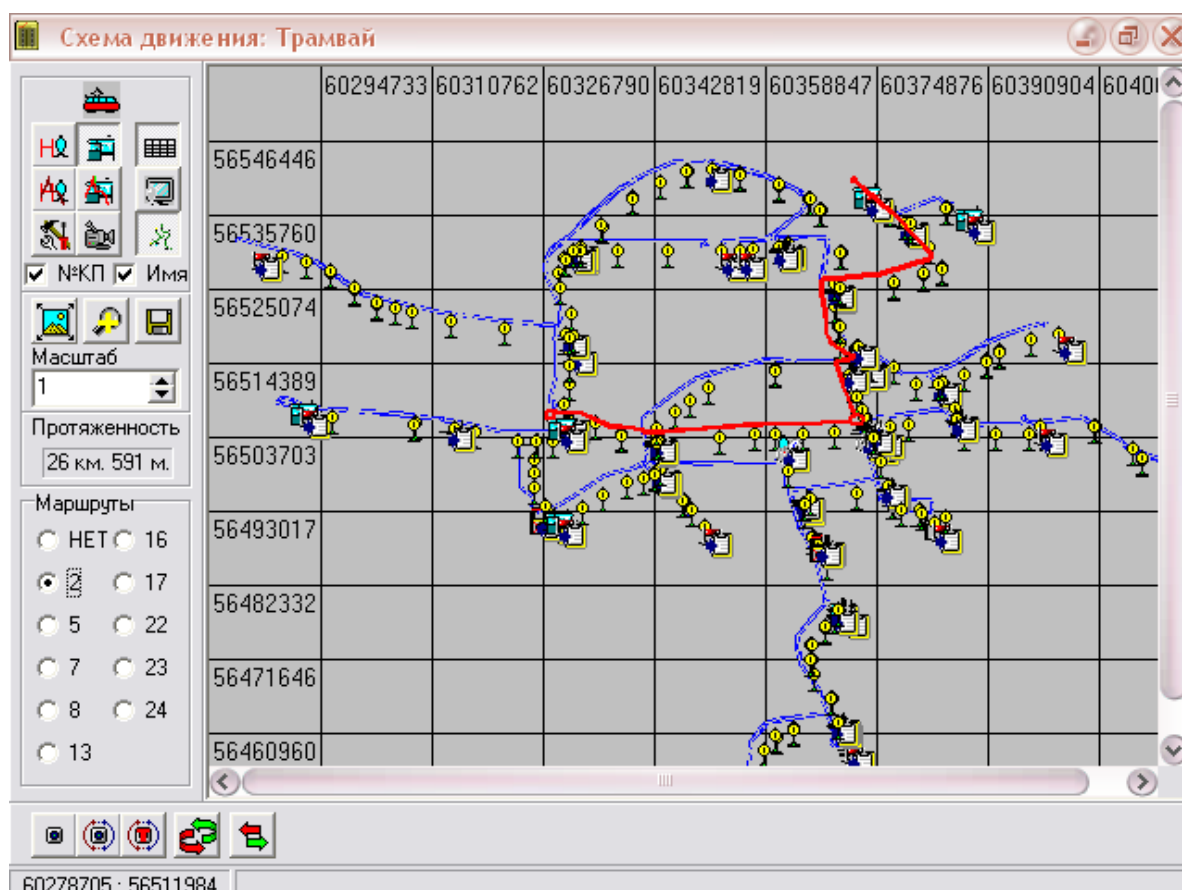


Рис. 14.5. Маршрутная схема движения трамвая

На основе составленного расписания автоматически рассчитываются плановые показатели (рис. 14.6) по каждой смене, по графикам, по маршрутам и в целом по депо.

Плановые показатели		
Маршрут	График	Смена
2	1	1
По сменам По графикам По маршрутам По депо		
Количество рейсов в 0-м направлении	3	
Количество рейсов в 1-м направлении	4	
Количество нулевых рейсов	1	
Количество оборотных рейсов	3.5	
Количество рейсов	3.80	
Поездо-километры в 0-м направлении	55км. 496м.	
Поездо-километры в 1-м направлении	0км. 0м.	
Поездо-километры нулевых рейсов	7км. 861м.	
Вагоно-километры нулевых рейсов	7км. 861м.	
Поездо-километры	103км. 476м.	
Количество рейсов в 0-м направлении для Disr	3	
Количество рейсов в 1-м направлении для Disr	4	
Длина нулевых выходов для Disr	7км. 861м.	
Длина нулевых заходов для Disr	0км. 0м.	
Количество спец-рейсов	0	
Длина спец-рейсов	0км. 0м.	
Начало работы ПЕ	06:11	
Остановка начала работы	ст. ВИЗ (№ 52)	
Окончание работы ПЕ	13:59	
Остановка окончания работы	ст. Фрезеровщиков (№ 9)	
Продолж. смены (включая обеда)	07:48	
Поездо-часы	07:33	
Вагоно-часы	07:33	
Начало работы водителя	05:51	
Окончание работы водителя	14:04	
Продолж. работы водителя	08:13	

Рис. 14.6. Расчет плановых показателей

Программный комплекс «Подготовка нарядов водителей и кондукторов» предназначен для составления и ведения нарядов на работу водителей и кондукторов, ПЕ маршрутизированного транспорта.

На рис. 14.7 представлена концептуальная модель «Подготовка нарядов водителей и кондукторов».

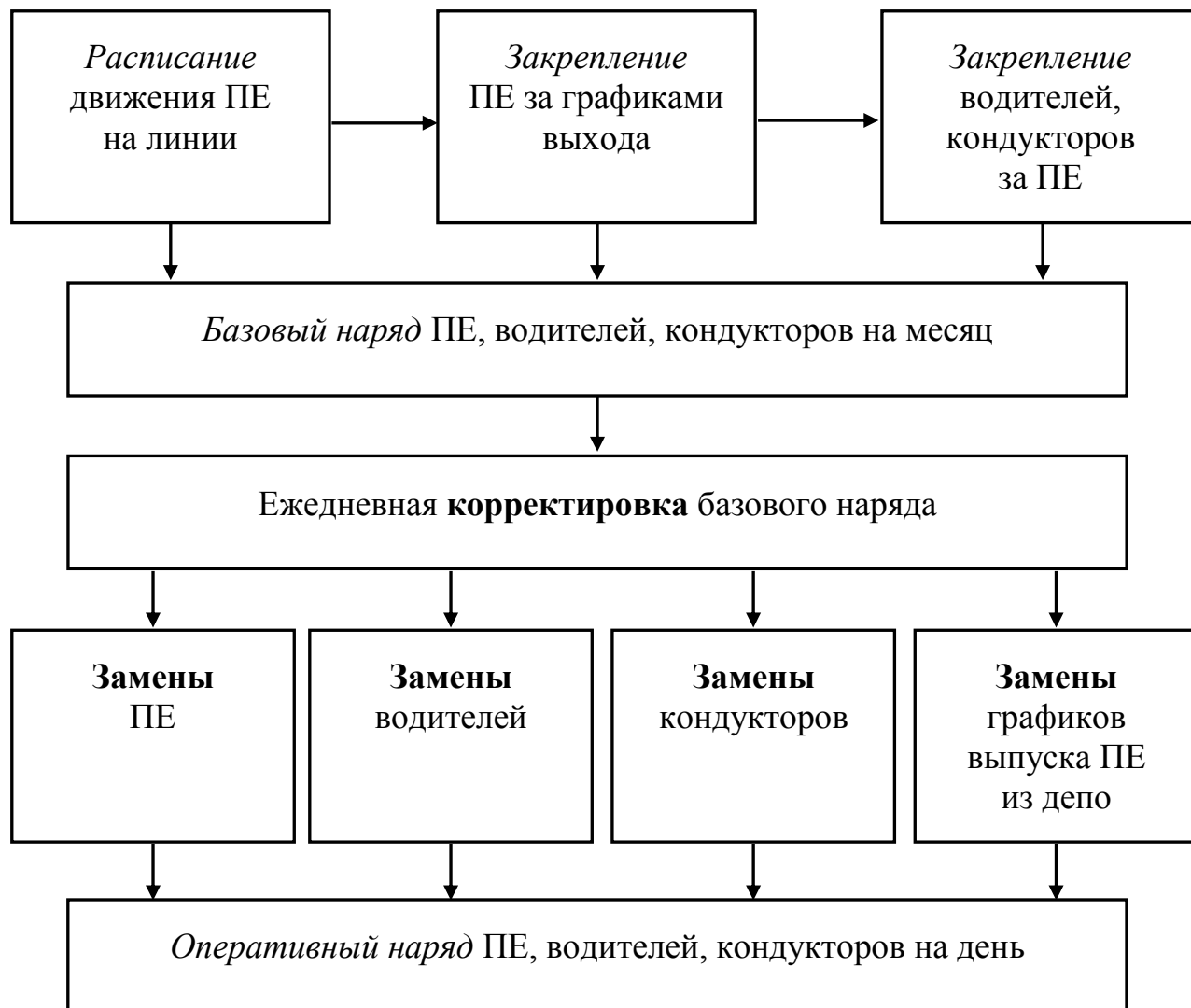


Рис.14.7. Концептуальная модель «Подготовка нарядов водителей и кондукторов»

Работа нарядчика транспортного предприятия включает в себя:

1) формирование наряда маршрута (плана работы водителя и кондуктора на месяц):

- группировка ПЕ;
- закрепление ПЕ за маршрутом и группой;
- закрепление водителя за ПЕ;
- закрепление смены за водителем (кондуктором);
- корректировка рабочего времени водителя (кондуктора) в соответствии с нормой;
- подготовка наряда на последующие дни для диспетчера выпуска с учетом внесенных изменений;

2) формирование справок и дополнительных отчетов по нарядам на работу водителей и кондукторов;

3) подготовка итоговых выходных данных для использования другими программами;

4) ежедневная корректировка наряда:

- замена ПЕ;
- замена водителя (кондуктора);
- замена смен;
- замена графиков выхода.

Исходными данными для формирования базовых нарядов являются расписание движения ПЕ, закрепление графиков за вагонами, закрепление номеров смен и табельных номеров водителей и кондукторов, Ф.И.О. водителей и кондукторов, изменения расстановки графиков, изменения закреплений табельных номеров и смен водителей, справочники дней осмотра ПЕ, режимов работы графиков, кодов отсутствия водителей, отработанное время работы водителей.

Программный комплекс «Подготовка нарядов водителей и кондукторов» позволяет автоматизировать следующие основные функции нарядчика:

- формирование основных групп ПЕ для проведения еженедельного технического осмотра каждой;
- закрепление смен на каждый день месяца за водителем (кондуктором);
- определение количества рабочих дней за месяц;
- подсчет ежедневного рабочего времени и отработанного за месяц;
- закрепление графика выхода за ПЕ с учетом позднего окончания второй смены;
- равномерное распределение ночных смен водителей (кондукторов);
- чередование графиков заступления водителей (кондукторов);
- чередование утреннего заступления водителей (кондукторов);
- анализ графика работы предыдущего дня;
- выравнивание заложенного по наряду рабочего времени водителей (кондукторов) за месяц и планового;
- закрепления за ПЕ водителей и кондукторов на текущий день;
- оперативный поиск водителя или кондуктора для замены отсутствующих.

В результате работы данного программного комплекса можно получить:

- базовый наряд ПЕ на маршрут;
- наряд водителя с учетом внесенных изменений;
- наряд кондуктора с учетом внесенных изменений;
- формы установленных отчетов, доступные в локальной (корпоративной) сети предприятия.

Программный комплекс «Диспетчер выпуска и движение подвижной единицы» предназначен для автоматизированного контроля над выпуском и движением ПЕ, работы водителя и кондуктора на данной ПЕ маршрутизированного городского транспорта.

Возможности программного комплекса:

- подготовка исходных данных по выпуску ПЕ на текущий день с учетом действующего расписания движения маршрутов транспорта;
- оперативное получение информации о закреплении ПЕ за маршрутом движения;
- корректировка базового наряда ПЕ по факту выхода на линию на текущее время;
- корректировка базового наряда водителя по факту выхода на линию на текущее время;
- корректировка базового наряда кондуктора по факту выхода на линию на текущее время;
- корректировка путевого листа водителя;
- автоматизация основных функций диспетчера выпуска ПЕ на линию:
 - изменение табельных номеров водителей и кондукторов;
 - введение замены ПЕ;
 - введение простоев ПЕ;
 - введение резерва водителей и кондукторов;
 - введение (удаление) нового графика маршрута.

Программный комплекс «Диспетчер выпуска и движение подвижной единицы» осуществляет контроль над выпуском на линию запланированного числа технически исправных ПЕ согласно расписанию выхода ПЕ и наряду водителей и кондукторов, а также производит своевременную оперативную замену ПЕ на технически исправные из резерва. Диспетчер выпуска следит за заступлением водителей в депо, оформляет путевой лист на ПЕ. Все изменения в путевом листе водителя и кондуктора или в базовом наряде ПЕ оперативно вносятся в данный программный комплекс.

На рис. 14.8 представлена концептуальная модель «Диспетчер выпуска и движение подвижной единицы».

Программный комплекс «Диспетчер выпуска и движение подвижной единицы» обеспечивает работу в двух основных режимах: «Оперативный наряд» и «Путевой лист».

В режиме «Оперативный наряд» (рис. 14.9) производится:

1. Ввод утренних замен ПЕ, табельных номеров водителей и кондукторов.
2. Изменение времени заступления кондуктора.
3. Ввод и удаление стажера-водителя.
4. Ввод и удаление нового кондуктора.
5. Ввод и удаление стажера-кондуктора.

В режиме «Путевой лист» (рис. 14.10) производится:

1. Ввод и удаление простоев ПЕ по технической неисправности, с последующей корректировкой любых введенных данных.
2. Ввод и удаление простоев водителей с учетом причин человеческого фактора, с последующей корректировкой любых введенных данных.
3. Ввод и удаление простоев кондукторов с учетом причин человеческого фактора с последующей корректировкой сопутствующих данных.



Рис. 14.8. Концептуальная модель «Диспетчер выпуска и движение подвижной единицы»

Диспетчер САС

Режим работы
☒ Нормальный

Основные данные
1 Маршрут
1 Гр. Вых.

Маршруты Г.В. ПЕ.

1	1 764
3	2 772
6	3 744
11	4 770
13	5 650
18	6 601
19	7 605

Дата: 15.04.2009
Дело: Западное
Конфигурация
Послать в "Где Т"

F2 - "ПУТЕВОЙ ЛИСТ"
Ctrl+R - Ввести "чистый" резерв
Ctrl+S - Добавить водителя стаж.
Ctrl+Ins - Добавить кондуктора
Shift+Ins - Добавить кондуктора стаж.
Ctrl+Del - Удалить кондуктора (стаж.)

Г.В.	См.	П.Е.1	П.Е.2	Т.Нач.	№ Вод.	Ф.И.О.	Т.Нач.	№ Кон.	Ф.И.О.
1	1	764	765	06:22	2296	ВЫРВИЧ Т.А.	06:22	9108	ИВАНЦОВА Н. В.
							06:22	9170	МИХАЙЛЮК Л. С.
1	2	764	765	14:16	2011	РОСТОВА О.А.	14:16	9008	МОРОЗОВА В. Г.
							14:16	9009	БУКИНА Л. А.
2	1	772	773	06:00	2075	ПОПОВА Л.Н.	06:00	9007	КОЛТЫШЕВА С. П.
							06:00	9142	
2	2	772	773	15:33	2279	КУЧЕРОВА И.В.	15:33	9132	КУЧУКОВА Р. А.
							15:33	9427	КЛЕВАКИНА Л. В.
3	1	744	745	06:10	2240	ХАМИДУЛЛИНА В.А.	06:10	9564	КОЛТАШОВА С. А.
							06:10	9577	ГОЛОВИНА Е. А.
3	2	744	745	14:09	2013	РЯБКОВА М.Б.	14:09	9564	КОЛТАШОВА С. А.
							14:09	9577	ГОЛОВИНА Е. А.

Рис. 14.9. Оперативный наряд

Диспетчер САС

Режим работы
☒ Нормальный

Основные данные
1 Маршрут
1 Гр. Вых.

Маршруты Г.В. ПЕ.

1	1 764
3	2 772
6	3 744
11	4 770
13	5 650
18	6 601
19	7 605

Дата: 15.04.2009
Дело: Западное
Конфигурация
Послать в "Где Т"

F2 - "ОПЕРАТИВНЫЙ НАРЯД"
Ctrl+P - Ввести простой водителя
Ctrl+K - Ввести простой кондуктора
Ctrl+T - Измен. время кондуктора
Ctrl+Z - Ввести замену
Ctrl+D - Удалить замену
Ctrl+G - Ввести график вне распис.
Ctrl+U - Удалить график

ПЕ № 764 765 **ПУТЕВОЙ ЛИСТ ЗА: 15.04.2009 г.** расп. от 01.04.2009 буд.

1 Смена

№ Вод.	Код Пр.	Код Уч.	Т.нач.	Т.кон.	№ Кон.	Код Пр.	Т.нач.	Т.кон.
2296	00	00	06:22	14:16	9108	00	06:22	14:16
					9170	00	06:22	14:16

НГВ 0 № п.о.: 7 Пл.=Ф.

№ Отм.	№ Мар.	Повт.	Оц.Р.

2,3 Смены

№ Вод.	Код Пр.	Код Уч.	Т.нач.	Т.кон.	№ Кон.	Код Пр.	Т.нач.	Т.кон.
2011	00	00	14:16	22:14	9008	00	14:16	22:14
					9009	00	14:16	22:14

НГВ 0 № п.о.: 7 Пл.=Ф.

№ Отм.	№ Мар.	Повт.	Оц.Р.

Рис. 14.10. Путевой лист

4. Ввод и удаление замены ПЕ в течение смены.
5. Ввод и удаление замены водителя в течение смены.
6. Изменение времени работы кондуктора.
7. Ввод и удаление нового графика, не отраженного в расписании.
8. Проведение оценки рейса по факту выполнения пробега.
9. Подготовка итоговых выходных данных для использования другими программными комплексами системы (время работы ПЕ, водителей и кондукторов на линии, простои ПЕ и водителей на линии, плановый и фактический пробег ПЕ, плановое и фактическое количество рейсов ПЕ на линии, плановая и фактическая скорость движения).

В результате работы данного программного комплекса можно получить:

- итоги работы водителей и кондукторов за день;
- показатели работы депо за день.

Программный комплекс «Табель учета рабочего времени водителей и кондукторов» позволяет вести электронный учет рабочего времени водителей и кондукторов, отработанного ими сверхустановленной нормы (снятие с выходного, сверхурочные) для текущей и последующей оплаты. Методика автоматизированного расчета рабочего времени работников, которым установлен суммированный учет рабочего времени, утверждена на предприятии ЕМУП ТТУ с учетом Трудового кодекса России.

Программный комплекс «Табель учета рабочего времени водителей и кондукторов» позволяет осуществлять:

- получение отчетных данных по каждому водителю и кондуктору;
- корректировку нормы рабочего времени на основании причин отсутствия водителя и/или кондуктора на основном рабочем месте;
- расчет рабочего времени в случае снятия с выходных дней;
- расчет сверхурочных часов рабочего времени.

На рис. 14.11 представлена концептуальная модель «Табель учета рабочего времени водителей и кондукторов».

Исходные данные программного комплекса «Табель учета рабочего времени водителей и кондукторов»:

- плановый индивидуальный наряд-закрепление по режимам работы водителей и кондукторов на месяц (квартал), предварительно рассчитанный и откорректированный в депо с учетом текущей работы и нормативно-справочных данных;
- индивидуальный производственный календарь на год с учетом праздничных дней и переносов рабочих и выходных дней;
- отработанное время (общее время, время по тарифам, праздничное время, время простоев);
- причины отсутствия водителей и кондукторов на основном рабочем месте;
- прочие явки на работу.

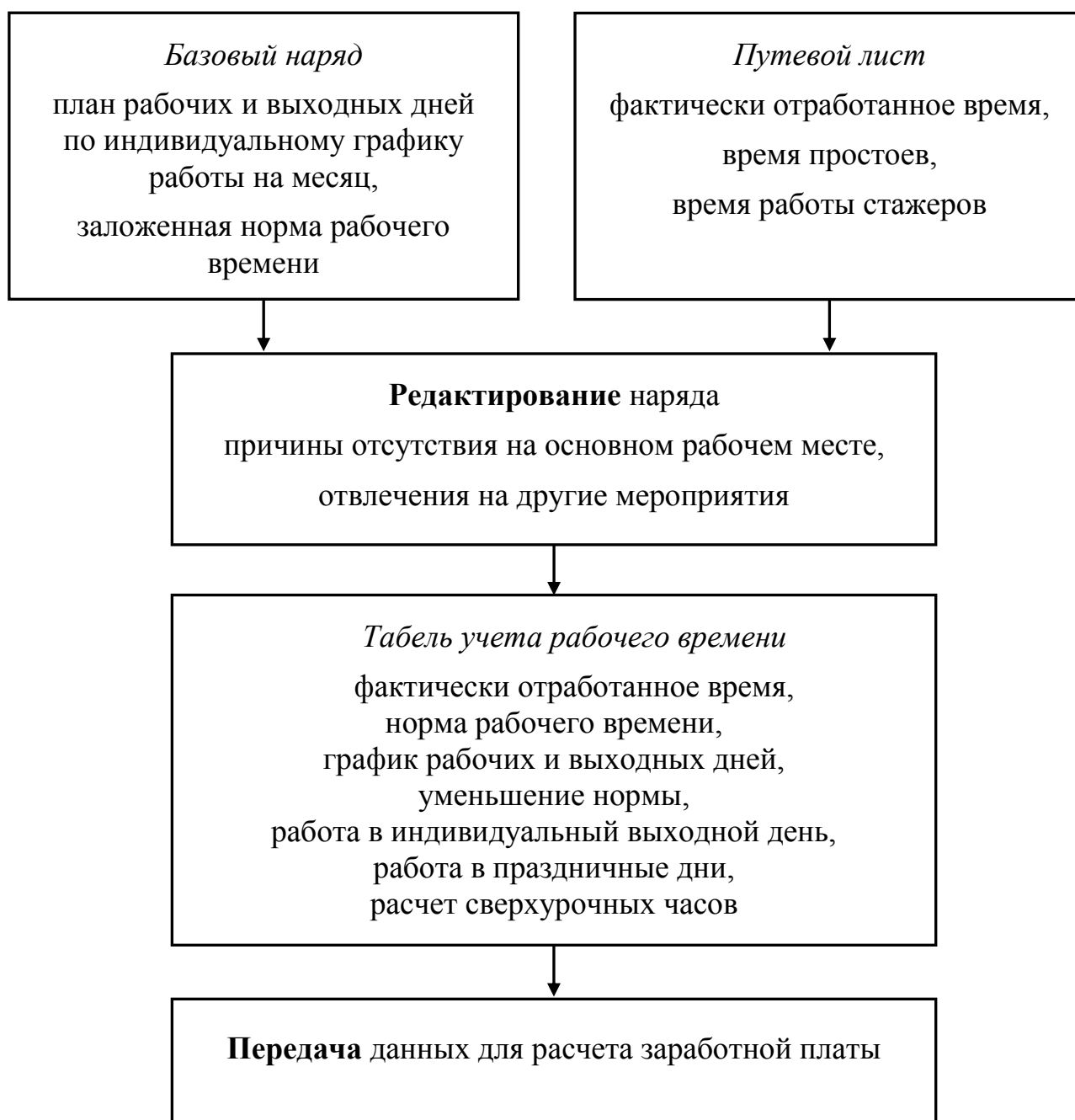


Рис. 14.11. Концептуальная модель «Табель учета рабочего времени водителей и кондукторов»

Общее время, время по тарифам, праздничное время, время простоев – это расчетные величины из программного комплекса «Путевой лист водителя и кондуктора». Причины отсутствий вводятся нарядчиком на основании соответствующих документов. Прочие явки – перевод на другую работу для выполнения производственных и иных работ, документально подтвержденных соответствующими документами.

В результате работы данного программного комплекса можно получить:

- табель учета рабочего времени водителей и кондукторов (рис. 14.12);
- данные для расчета заработной платы водителей и кондукторов.

- экономия или перерасход топлива;
- остатки горючего на начало текущего и следующего месяца;
- машино-часы в движении.
- отчетные документы о работе автотранспортной службы (АТС), доступные на Web-сервере.

Кроме того, при наличии данных о заправке водителей на автозаправочных станциях можно провести сравнительный анализ данных путевого листа автомашины и данных с нефтезаправочных станций о заправке бензина на каждую автомашину.



Рис. 14.13. Концептуальная модель «Путевой лист автотранспортной службы»

В рамках городской целевой программы «Электронный Екатеринбург», утвержденной Екатеринбургской городской думой, внедряется программный комплекс «Транспорт города Екатеринбурга» в виде информационного сайта. Организатором является администрация города Екатеринбурга в лице Комитета по промышленности, науке, связи и информационным технологиям и Комитета по транспорту и организации дорожного движения. В настоящее время выполнена интеграция сайта с информационно-коммуникационной системой службы автоматизации и связи ЕМУП ТТУ. Информационный сайт находится в открытом доступе и предназначен для всех пользователей Интернет-сети.

В программном комплексе «Транспорт города Екатеринбурга» формируются электронные документы в формате *html* по утвержденной в ЕМУП ТТУ структуре, содержащие следующую информацию:

- маршруты движения трамваев и троллейбусов в г. Екатеринбург;
- конфигурация маршрутов трамваев и троллейбусов в г. Екатеринбург;
- расписание движения трамваев и троллейбусов в г. Екатеринбург для населения;
- выполнение плана рейсов трамваев и троллейбусов;
- выпуск подвижного состава трамваев и троллейбусов по времени;
- справка о сходе трамваев и троллейбусов;
- интервал движения по маршрутам по периодам времени трамваев и троллейбусов;
- сведения о движении поездов на линии по маршрутам трамваев и троллейбусов;
- время прохождения станции трамваев и троллейбусов;
- расписание (режим) работы трамваев и троллейбусов;
- задержки движения на линии и в депо по техническим неисправностям трамваев и троллейбусов;
- показатели работы ЕМУП ТТУ (рис. 14.14).

Информационный сайт «Транспорт города Екатеринбурга» позволяет:

- информировать население о маршрутах трамваев и троллейбусов;
- получить информацию о маршрутах следования любого выбранного маршрута, т.е. описание порядка следования всех остановок маршрута;
- получить сведения об интервале движения маршрутов по конечным станциям по периодам времени;
- получить расписание движения трамваев и троллейбусов по выбранной пользователем станции или остановки города Екатеринбурга.

На основе данных информационного сайта «Транспорт города Екатеринбурга» Комитет по транспорту и организации дорожного движения может:

- осуществлять ежедневный контроль выполнения основных показателей ЕМУП ТТУ;
- получать итоговые данные выполнения показателей работы ЕМУП ТТУ за период (за любой день, за любой месяц, за любой год);
- получать ежедневный контроль за регулярностью движения по маршрутам

трамваев и троллейбусов, а также итоговые данные за прошлый день, за прошлый месяц;

- знать о выпуске подвижного состава на улицы города в любой заданный момент.

Показатели работы по ЕМУП "ТТУ" - Microsoft Internet Explorer

Адрес: http://root/emup2/index.php

**Показатели работы по
ЕМУП «Трамвайно-троллейбусное управление»
за февраль 2007.**

Депо	Вагоночасы		Рейсы					Регу- ляр- ность	Пробег (тыс. км)	
	план (расп.)	факт	план (расп.)	факт	граф	сорв.	Целосы- постен		план (расп.)	факт
Северное (№ 17)	24779.98	24295.32	7441.36	7021.82	6875.40	419.54	78.75	92.4	362.654	346.663
Южное (№ 17)	21873.97	21655.33	7220.72	6592.79	6500.55	627.93		90.0	301.410	274.520
Западное (№ 17)	24067.59	23615.90	10079.60	9349.55	9243.82	730.04	7.54	91.7	325.862	301.856
Октябрьское (№ 17)	19500.16	19285.56	11661.47	9266.34	9126.35	2395.13	25.41	78.3	299.175	238.234
Орджоникидзевское (№ 17)	19451.82	18146.74	21470.99	18194.73	18135.77	3276.26	648.80	84.5	289.595	227.191
Трамвай	83808.71	82278.52	29338.97	27443.74	27099.35	1895.23	123.07	92.4	1174.025	1103.336
Троллейбус	45664.18	43779.21	38914.69	33060.08	32861.14	5854.61	823.21	84.4	691.240	564.838
Итого по МУП ЕТТУ	129472.89	126057.73	68253.67	60503.82	59960.49	7749.84	946.28	87.8	1865.264	1668.174

Начальник службы
автоматики и связи

Дроздов Н.П.

Рис. 14.14. Показатели работы ЕМУП ТТУ

Все данные расписания выдаются за конкретно выбранный день с учетом рабочих, выходных и праздничных дней, как календарных, так и перенесенных рабочих и выходных дней. Оперативность данных позволяет наиболее полно информировать население об изменениях в движении трамваев и троллейбусов, как постоянных (например, ввод нового маршрута или открытие новых участков движения), так и временных (во время закрытий определенных участков движения для производства ремонтных работ).

По результатам мониторинга дохода по перевозке пассажиров в виде отчетов информационно-коммуникационной системы при большой загрузке маршрута может возникнуть необходимость в добавлении подвижных единиц и тем самым уменьшении интервала между ними, например в час-пик, или даже в добавлении нового маршрута на загруженном участке города.

Рассмотрим пример. В Южном трамвайном депо г. Екатеринбурга был введен маршрут № 9 в связи с большой загрузкой данного участка города, который работал по системе: рабочие дни только в час-пик. Хотя, рассматривая общий доход (рис. 14.15), видно, что маршрут № 9 не является доходным, так как общая выручка на маршруте не велика. Однако, анализируя работу маршрута по величине дохода на один график маршрута (рис. 14.16), видно, что маршрут № 9 является прибыльным (5793,67 руб. на график). Поэтому было принято решение выпускать на линию этот маршрут и в выходные дни, причем

доходность маршрута по выходным дням тоже оказалась велика. Анализ дохода по графикам маршрута на 1 машино-час позволяет определить малодоходный график, который можно корректировать по времени выхода из депо (увеличить или уменьшить время работы), например в выходные дни.

Выручка по маршрутам за день и по число - Microsoft Internet Explorer

13.12.2010 12:23:21
Депо Южное

Выручка по маршрутам по 4.12.2010

М-т	Билеты		Проездные		Терминал	
	за 4.12	по 4.12	за 4.12	по 4.12	за 4.12	по 4.12
4	85320.00	452268.00		53480.00	2832.00	14366.00
9	24246.00	149634.00		20495.00	700.00	3923.00
10	121788.00	528966.00	5120.00	105400.00	3815.00	15723.00
14	71316.00	454554.00	2510.00	71450.00	1955.00	12769.00
15	207900.00	985068.00		116050.00	5936.00	27089.00
20	39330.00	180540.00	840.00	37280.00	1923.00	7376.00
25	103824.00	507222.00	1320.00	66315.00	3066.00	14865.00
27	72414.00	367686.00		52980.00	2099.00	10599.00
31	20304.00	163620.00		19995.00	896.00	6056.00
Итого по депо	746442.00	3789558.00	9790.00	543445.00	23222.00	112766.00

Готово

Рис. 14.15. Выручка по маршрутам и депо

Выручка по рабочим сменам - Microsoft Internet Explorer

13.12.2010 12:26:25
Депо Южное

Рабочая смена за 01.12.2010

М-т	1 смена			2 смена			12-часовые			Перерывные			Итого	
	Кол-во ПЕ	Сумма	Терминал, (кол-во)	Кол-во ПЕ	Сумма	Терминал, (кол-во)	Кол-во ПЕ	Сумма	Терминал, (кол-во)	Кол-во ПЕ	Сумма	Терминал, (кол-во)	Сумма	Терминал, (кол-во)
4	12	63936	2007	10	60138	1359							124074	3366
9										8	39870	922	39870	922
10	14	81072	2813	14	66312	1072							147384	3885
14	15	67842	2590	13	58032	627							125874	3217
15	28	128124	4833	27	134280	1471							262404	6304
20	5	23976	1435	5	23778	287							47754	1722
25	15	76752	2467	16	64242	1260							140994	3727
27	7	44730	1322	7	41328	894				3	19008	525	105066	2741
31	4	19314	1031	4	27540	561							46854	1592
Итого		505746	18498		475650	7531					58878	1447	1040274	27476

Начальник отдела сборов

Готово

Рис. 14.16. Выручка по рабочим сменам в рабочий день

Однако иногда существование малодоходных маршрутов необходимо как городской заказ на перевозку пассажиров из районов города, в которых

отсутствует альтернативный транспорт. Например, маршрут № 11 «Зеленый остров – ВИЗ», или в связи с ремонтом дороги по ул. Московская временно отсутствует другой вид общественного транспорта, поэтому был введен маршрут № 33 «ЦПКиО – Дворец спорта».

На рис. 14.17 показан пример отчета о перевозках пассажиров всего предприятия, а более подробный отчет по перевозкам по маршрутам за период показан на рис. 14.18.

http://192.168.0.82/depo/bue/tranzpodrNgrM.php?NAD=21&NAM=12&NAG=2010&TEKDEN1=1&TEKDEN2=1&TEKM - Microsoft Internet Explorer

Файл Правка Вид Избранное Сервис Справка

21.12.2010 16:16:57

Отчет об объеме транспортных перевозок по данным ОАО "И-сеть" за период с 01.12.2010 по 01.12.2010 по состоянию на 21.12.2010

Дено (Класс в этой колонке откро- ет подроб- ности по марш- рутам)	Льготники		Школьники		Студенты		Об. пассаж								Итого		Всего с учетом суммы за проезд по проездным с оплаткой за поездку и банк.картам		Нал. расчет "Воз- врат (для справ- ки)
	Итого льгот						Проезды Итого		Проезды Итого		Проезды Итого		Проезды Итого		Проезды Итого				
	Проезды Итого	Эл.кош.	Проезды Итого	Эл.кош.	Проезды Итого	Эл.кош.	Проезды Итого	Эл.кош.	Проезды Итого	Эл.кош.	Проезды Итого	Эл.кош.	Проезды Итого	Эл.кош.	Проезды Итого	Эл.кош.			
	Сумма (К-во)	Сумма (К-во)	Сумма (К-во)	Сумма (К-во)	Сумма (К-во)	Сумма (К-во)	Сумма (К-во)	Сумма (К-во)	Сумма (К-во)	Сумма (К-во)	Сумма (К-во)	Сумма (К-во)	Сумма (К-во)	Сумма (К-во)	Сумма (К-во)	Сумма (К-во)			
	40-42 /50-52,60	46-47 /55-57	53,35 /36,39	58	54,30 /31,34	59	21,25,26 /70	80-82 /85	77,74 /72		61,65		11-15 /20,22		14945 /14875		22338.00 (1710)		
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18		
Окт	13099 /13099	16884.00 (1407)	90 /80	18.00 (1)	268 /256	0 (0)	323 /308	33 /0	1040 /961	15020.58 /13723.08	1656.00 (92)	5436.00 (302)	14945 /14875	22338.00 (1710)	16655	39014.58	54.00 (3)		
	./0	./0	./0	./0	./12	./13	./13	./23	./11	./192.50	./0	./0	./56	./14	./0	./0	./0		
Орл	25771 /25771	29712.00 (2476)	296 /255	54.00 (3)	1028 /1024	0 (0)	1145 /1099	49 /0	2638 /2483	38000.99 /35457.24	2826.00 (157)	17172.00 (954)	31084 /30944	46938.00 (3433)	34517	87764.99	18.00 (1)		
	./0	./0	./28	./13	./4	./18	./41	./0	./135	./2193.75	./0	./0	./91	./0	./0	./0	./0		
Сев	23741 /23741	31644.00 (2637)	137 /114	0 (0)	736 /696	0 (0)	799 /690	62 /0	3187 /2964	45988.42 /42325.92	3294.00 (183)	15533.82 (863)	28845 /28611	47177.82 (3500)	32345	96460.24	126.00 (7)		
	./0	./0	./21	./0	./40	./0	./96	./55	./192	./5120.00	./0	./0	./212	./0	./0	./0	./0		
Южн	19853 /19853	24516.00 (2043)	157 /111	18.00 (1)	820 /761	0 (0)	899 /812	83 /0	2710 /2546	39055.63 /36356.88	3384.00 (188)	12977.82 (721)	24710 /24435	37511.82 (2765)	27475	79951.45	18.00 (1)		
	./0	./0	./41	./0	./59	./0	./78	./76	./137	./2226.25	./0	./0	./254	./0	./0	./0	./0		
Зап	28997 /28997	37332.00 (3111)	230 /171	54.00 (3)	943 /899	18.00 (1)	1018 /892	120 /0	3327 /3096	48003.38 /44210.88	4338.00 (241)	17496.00 (972)	34876 /34527	54900.00 (4087)	38963	107241.38	18.00 (1)		
	./0	./0	./57	./0	./84	./0	./118	./113	./200	./5320.00	./0	./0	./332	./0	./0	./0	./0		
Тран- зай	72591 /72591	93492.00 (7791)	524 /396	72.00 (4)	2499 /2356	18.00 (1)	2716 /2394	265 /0	9224 /8606	133047.43 /122893.68	11016.00 (612)	46007.64 (2556)	88431 /87573	139589.64 (10352)	98783	283653.07	162.00 (9)		
	./0	./0	./119	./0	./143	./0	./292	./244	./529	./8596.25	./0	./0	./798	./0	./0	./0	./0		
	./0	./0	./0	./0	./0	./0	./0	./0	./89	./1557.5	./0	./0	./0	./0	./0	./0	./0		

Готово

Местная интрасеть

Рис. 14.17. Отчет об объеме транспортных перевозок предприятия

Перевозка пассажиров по часам суток отражена на рис. 14.19. Анализ перевозок по часам суток помогает анализировать наполняемость транспорта, выявлять «пиковые» часы или равномерную перевозку в течение суток на каком-либо маршруте. Примерная перевозка пассажиров на маршруте № 9 за сутки: 922 пассажиров делим на 6 графиков, получаем ≈ 153 пассажира на график в среднем, а на маршруте № 4 в сутки 10 графиков: $3366 / 10 \approx 336$ пассажира на график в среднем. Но маршрут № 9 выходит на линию только в утреннее и вечернее время, и сравнивать эти параметры не корректно. Сравним эти же маршруты в период 7:00-8:00 – маршрут № 4 ($278 / 10 \approx 27$ пассажиров на график в среднем) и маршрут № 9 ($154 / 6 \approx 25$ пассажиров на график в среднем). И мы видим, что объем перевозки на данных маршрутах в один период времени одинаковый.

Рис. 14.18. Отчет об объеме транспортных перевозок структурного подразделения

Рис. 14.19. Перевозка пассажиров по часам суток

181

пассажир должен быть уверен, что он сможет воспользоваться городским электротранспортом и в утренние и ночные часы.

Сравним этот же период у маршрута № 11 ($59 / 2 \approx 28$ пассажиров на график в среднем) и у маршрута № 33 ($100 / 4 \approx 25$ пассажиров на график в среднем). Опять получаем примерно то же количество пассажиров в среднем (от 25 до 30). Самым наполняемым маршрутом по этому показателю получается маршрут № 1 ($724 / 12 \approx 60$ пассажиров на график в среднем, рис. 14.20). Вывод: ввод учета пассажиров по Е-карте помогает предприятию вести более точный учет перевозки пассажиров, позволяет более точно анализировать наполняемость подвижных единиц по часам суток. Это позволяет оптимизировать расписание движения, в частности ввести новые графики на отдельные маршруты, ввести сдвоенные поезда (сплотки), изменить графики работы в утреннее, вечернее и ночное время. Учет пассажиропотока на городском общественном транспорте и общий анализ о перевезенных пассажирах позволяет решать задачу адаптации городского электротранспорта к потребностям пассажиров города.

http://192.168.0.82/depo/buz/tranzpotok.php?ТЕКТР=303&ТЕКДЕН1=1&ТЕКДЕН2=1&ТЕКМЕSEC=1&ТЕКГОД=20 - Microsoft Internet Explorer

ФайлПравкаВидИзбранноеСервисСправка

Тип расписания	Маршрут	Количество пассажиров по часам суток в процентах от общего кол-ва (кол. 24)																								позд 24: 00	Итого
		ран 05: 00	05: 00- 06:00	06: 00- 07:00	07: 00- 08:00	08: 00- 09:00	09: 00- 10:00	10: 00- 11:00	11: 00- 12:00	12: 00- 13:00	13: 00- 14:00	14: 00- 15:00	15: 00- 16:00	16: 00- 17:00	17: 00- 18:00	18: 00- 19:00	19: 00- 20:00	20: 00- 21:00	21: 00- 22:00	22: 00- 23:00	23: 00- 24:00						
			00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00			
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24				
будни	1	0	8	140	724	686	417	569	622	584	348	395	478	446	632	424	270	190	39	6	0	0	0	6978			
		0	0.1	2.0	10.4	9.8	6.0	8.2	8.9	8.4	5.0	5.7	6.9	6.4	9.1	6.1	3.9	2.7	0.6	0.1	0	0	0	100.0			
будни	3	0	27	134	280	215	249	206	302	207	237	220	168	251	229	157	130	49	16	2	0	0	0	3079			
		0	0.9	4.4	9.1	7.0	8.1	6.7	9.8	6.7	7.7	7.1	5.5	8.2	7.4	5.1	4.2	1.6	0.5	0.1	0	0	0	100.0			
будни	6	0	12	214	427	204	165	371	370	461	506	390	422	363	424	277	253	117	59	5	0	0	0	5040			
		0	0.2	4.2	8.5	4.0	3.3	7.4	7.3	9.1	10.0	7.7	8.4	7.2	8.4	5.5	5.0	2.3	1.2	0.1	0	0	0	100.0			
будни	11	0	4	12	59	52	27	38	42	37	30	49	25	41	62	25	21	14	5	1	0	0	0	544			
		0	0.7	2.2	10.8	9.6	5.0	7.0	7.7	6.8	5.5	9.0	4.6	7.5	11.4	4.6	3.9	2.6	0.9	0.2	0	0	0	100.0			
будни	13	0	2	45	225	144	103	66	201	275	245	231	201	229	284	121	173	50	14	2	0	0	0	2611			
		0	0.1	1.7	8.6	5.5	3.9	2.5	7.7	10.5	9.4	8.8	7.7	8.8	10.9	4.6	6.6	1.9	0.5	0.1	0	0	0	100.0			
будни	18	0	9	65	337	249	274	244	306	216	332	246	251	338	281	268	133	129	23	6	0	0	0	3707			
		0	0.2	1.8	9.1	6.7	7.4	6.6	8.3	5.8	9.0	6.6	6.8	9.1	7.6	7.2	3.6	3.5	0.6	0.2	0	0	0	100.0			
будни	19	0	5	194	559	447	358	435	493	424	289	254	400	493	569	324	223	127	39	22	0	0	0	5655			
		0	0.1	3.4	9.9	7.9	6.3	7.7	8.7	7.5	5.1	4.5	7.1	8.7	10.1	5.7	3.9	2.2	0.7	0.4	0	0	0	100.0			
будни	21	0	35	122	272	228	214	157	161	210	194	187	286	275	267	213	104	64	44	11	1	0	0	3045			
		0	1.1	4.0	8.9	7.5	7.0	5.2	5.3	6.9	6.4	6.1	9.4	9.0	8.8	7.0	3.4	2.1	1.4	0.4	0.0	0	0	100.0			
будни	26	0	4	67	226	239	222	287	305	271	239	301	310	277	423	304	170	144	45	7	0	0	0	3841			
		0	0.1	1.7	5.9	6.2	5.8	7.5	7.9	7.1	6.2	7.8	8.1	7.2	11.0	7.9	4.4	3.7	1.2	0.2	0	0	0	100.0			
будни	32	0	1	74	238	234	213	175	224	237	231	234	246	273	304	200	156	159	79	12	0	0	0	3290			
		0	0.0	2.2	7.2	7.1	6.5	5.3	6.8	7.2	7.0	7.1	7.5	8.3	9.2	6.1	4.7	4.8	2.4	0.4	0	0	0	100.0			
будни	33	0	5	47	100	101	77	73	80	90	96	38	108	57	80	83	101	29	9	0	0	0	0	1174			
		0	0.4	4.0	8.5	8.6	6.6	6.2	6.8	7.7	8.2	3.2	9.2	4.9	6.8	7.1	8.6	2.5	0.8	0	0	0	0	100.0			

8:33

ГотовоМестная интрасеть

Рис. 14.20. Перевозка по часам суток Западного депо

На предприятии ведется сравнение работы структурных подразделений – депо, например по доходу от перевозки пассажиров (рис. 14.21). Это стимулирует работу подразделений. Низкие показатели сбора денежных средств в Октябрьском депо объясняются общей схемой городского транспорта г. Екатеринбурга. Все маршруты Октябрьского депо проходят по загруженным городским магистралям, параллельно идет большое число коммерческого пассажирского транспорта, что, несомненно, уменьшает долю депо в общей

работе предприятия. Заказ на перевозку пассажиров распределяется между депо, и учет пассажиров по Е-карте помогает планировать и распределять работу между депо, в первую очередь изменять расписание движения.

На предприятии ведется особый контроль и учет за перевозкой льготных категорий пассажиров (рис. 14.22).

21.12.2010 16:19:37

Отчет по данным отдела сборов ЕМУП ТТУ за период с 01.12.2010 по 01.12.2010

Депо	Маршрут	Таб. номер	Кол-во операций	Эл. кошелек сумма	Проездной кол-во	Наличная оплата сумма	Возврат кол-во	Доля депо
Октябрьское троллейбусное депо			16658	22338	14945	54	0	11.11
Орджоникидзевское троллейбусное депо			34519	46938	31084	18	1	23.02
Северное трамвайное депо			32352	47178	28845	126	0	21.57
Южное трамвайное депо			27476	37512	24710	36	0	18.32
Западное трамвайное депо			38964	54900	34876	18	0	25.98
Всего по ТТУ			149969	208866	134460	252	1	100

16:20

Готово Местная интрансеть

Рис. 14.21. Итоговые данные

http://192.168.0.82/depo/bup/tranzpodrNNL.php?NAD=21BNAM-12BNAG-2010TEKEDEN1-18TEKDENZ-18TEKME5 - Microsoft Internet Explorer

21.12.2010 16:21:08

Только льготные категории
Отчет об объеме транспортных перевозок по данным ОАО "И-сеть"
за период с 01.12.2010 по 01.12.2010
по состоянию на 21.12.2010

Депо	Льготники								Школьники	Студенты	Итого		Всего	
	Федеральн		Регион		Муниц		Итого льгот				Проезды Итого /все виды /трам /трол /трам-трол	Проезды Итого /все виды /трам /трол /трам-трол		
	Проезды Итого /все виды /трам /трол /трам-трол	Эл.кош. Сумма (К-во)	Проезды Итого /все виды /трам /трол /трам-трол	Эл.кош. Сумма (К-во)	Проезды Итого /все виды /трам /трол /трам-трол	Эл.кош. Сумма (К-во)	Проезды Итого /все виды /трам /трол /трам-трол	Эл.кош. Сумма (К-во)						
	40,50,60	45,55	41,51	46,56	42,52	47,57	40-42 50-52 60	45-47 55-57	53	54				
Окт	2061 /2061 /0 /0	3588.00 (299)	6027 /6027 /0 /0	7980.00 (665)	5011 /5011 /0 /0	5316.00 (443)	13099 /13099 /0 /0	16884.00 (1407)	90 /80 /0 /8 /2	268 /256 /0 /12 /0	13457 /13435 /0 /20 /2	16884.00 (1407)	14864	16884.00
Орд	4689 /4689 /0 /0	7752.00 (646)	11141 /11141 /0 /0	13116.00 (1093)	9941 /9941 /0 /0	8844.00 (737)	25771 /25771 /0 /0	29712.00 (2476)	296 /255 /0 /28 /13	1028 /1024 /0 /4 /0	27095 /27050 /0 /32 /13	29712.00 (2476)	29571	29712.00
Сев	4063 /4063 /0 /0	7704.00 (642)	9487 /9487 /0 /0	12528.00 (1044)	10191 /10191 /0 /0	11412.00 (951)	23741 /23741 /0 /0	31644.00 (2637)	137 /114 /21 /0 /2	736 /696 /40 /0 /0	24614 /24551 /61 /2	31644.00 (2637)	27251	31644.00
Итого	2227	5522.00	2197	10920.00	9419	8064.00	10952	24516.00	157	920	20920	24516.00	22973	24516.00

16:21

Готово

Местная интрансеть

Рис. 14.22. Учет льготных категорий пассажиров

Для удобства и привлечения пассажиров на городской электротранспорт разработана серия льготных видов оплаты проезда с помощью Е-карты:

- персональные карты для школьников, студентов и пенсионеров на 4 вида транспорта без ограничения количества поездок;
- персональные карты для школьников, студентов и пенсионеров на 4 вида транспорта с ограничением количества поездок по льготной цене за одну поездку;
- неперсональные карты для граждан на 1, 2 и 4 вида транспорта с ограничением количества поездок с льготной ценой за одну поездку.

Это позволяет заинтересовывать пассажиров пользоваться городским общественным транспортом и является одной из основных задач логистики – увеличение пассажиропотока, т.е. повышение конкурентоспособности предприятия в условиях жесткой конкуренции на рынке. Эти категории пассажиров имеют льготную цену за перевозку, и учет пассажиропотока позволяет точно знать количество перевезенных пассажиров и цену поездки каждой категории перевезенных пассажиров.

Таким образом, исполнение концепции логистики позволяет:

- гармонизировать внутренние бизнес-процессы предприятия, повысить эффективность деятельности функциональных подразделений;
- сократить затраты и уменьшить себестоимость продукции;
- увеличить количество потенциальных и фактических потребителей, удержать и расширить рынок услуг;
- повысить качество обслуживания заказов потребителей, укрепить репутацию и авторитет предприятия;
- повысить конкурентоспособность предприятия в условиях жесткой конкуренции на рынке.

14.2. УПРАВЛЕНИЕ И АНАЛИЗ РАБОТЫ ТРАНСПОРТНОЙ СЕТИ С ПОМОЩЬЮ СПУТНИКОВОЙ НАВИГАЦИИ

Сегодня резко повышены требования к качеству услуг сервиса пассажироперевозок. Для улучшения сервиса услуг по перевозке пассажиров с возможностью адаптации в условиях неопределенности окружающей среды, изучаются основные потоки движения пассажиров, а именно резкие колебания качественных и количественных характеристик, и таким образом организовываются графики движения, чтобы максимально ускорить перевозки пассажиров в различные точки города.

Оптимально составленное маршрутное расписание должно обеспечить: наименьшее время ожидания пассажиров и поездки к месту назначения; равномерное наполнение по всем перегонам маршрута; высокую регулярность на протяжении всего периода движения; высокую скорость сообщения при соблюдении безопасности поездок; эффективное использование транспорта, соблюдение режима труда и отдыха водителей и кондукторов; соблюдение плановых технических осмотров подвижного состава; согласованность интервалов движения по отправлению на узловых остановках; выполнение

плановых экономических показателей работы транспортного предприятия. В связи со значительными колебаниями пассажиропотоков по времени года и дням недели расписание движения составляют на весенне-летний и осенне-зимний периоды года, а также отдельно для рабочих, субботних и воскресных дней. Особое внимание уделяется определению необходимого числа рейсов в часы пик с учетом комфортной наполняемости подвижного состава городского электротранспорта для пассажиров.

Продукция транспортного предприятия – это качественная (с точки зрения бесперебойности, безопасности, ритмичности, всестороннего удовлетворения нужд и запросов населения) перевозка пассажиров, а следовательно, и эффективный доход, полученный от перевозки пассажиров.

С введением в г. Екатеринбурге с 2010 года электронного учета с помощью Е-карты на предприятии был налажен более точный учет перевозки пассажиров различных категорий: школьники, студенты, социальные пассажиры, обычные пассажиры. Это позволяет более точно анализировать наполняемость подвижных единиц по часам суток и, таким образом, оптимизировать расписание движения, в частности вводить новые графики на отдельные маршруты, сдвоенные поезда (сплотки), изменять графики работы в утреннее, вечернее и ночное время. Учет пассажиропотока и общий анализ о перевезенных пассажирах позволяет решать задачу адаптации городского электротранспорта к потребностям пассажиров города.

На рис. 14.23 представлена концептуальная модель анализа технико-экономических показателей городского транспорта на основе информационно-коммуникационной системы МУП ТТУ г. Екатеринбурга.

Анализ ритмичности работы транспортного предприятия включает в себя анализ различных справок и отчетов технико-экономического характера в разрезе работы подвижной единицы (ПЕ), водителя, маршрутов, депо и в целом работы всего предприятия. Пробег ПЕ на линии напрямую зависит от заложенного плана пробега и от качества работы на линии (бесперебойности, т.е. минимального количества простоев). В процессе анализа выполнения планов может быть дана и критическая оценка самого плана (расписания движения транспорта), его обоснованности с учетом запросов населения с целью повышения экономической эффективности и доходности.

Показатели работы транспортной сети делятся на 3 категории: пробеги, доходность, простои. Эти показатели формируются для ПЕ, маршрута, графика выхода ПЕ из депо, а также для каждого депо и в целом для всего предприятия в течение определенных временных промежутков (смена, сутки, неделя, месяц, год).

Основой определения плановых показателей работы является расписание маршрутизированного транспорта, по которому оформляется выпуск маршрутов, графиков подвижной единицы. На основе графиков движения создаются наряды подвижной единицы, водителей, кондукторов. Далее по факту наличия замен и простоев диспетчер выпуска и движения подвижной единицы формирует оперативный наряд и путевой лист, которые периодически редактируются. Далее формируется табель учета рабочего времени водителей и

кондукторов, по которому производится расчет зарплаты водителей и кондукторов. Параллельно определяются необходимые технико-экономические показатели. Например, наличие ПЕ на линии по времени суток, простои ПЕ на линии, в депо, отклонение от фактического выпуска ПЕ, регулярность движения маршрутов, суммарные показатели работы водителей, кондукторов и т.п.



Рис. 14.23. Концептуальная модель анализа технико-экономических показателей городского транспорта

Для автоматического сбора информации на линии транспортной сети до недавнего времени существовали только контрольные пункты, на которых снимались отметки о прохождении ПЕ. Данные поступали диспетчеру, который проверял отклонения движения ПЕ от расписания. В случае ДТП или других

задержек и простоев водитель по радиосвязи передавал эту информацию диспетчеру, который регулировал схему движения других ПЕ также по радиосвязи. Оперативность такого управления была пониженной.

В настоящее время оперативное управление работой транспортной сети в ЕМУП ТТУ осуществляется с помощью автоматизированной системы управления транспортом (АСУ-Т) на базе современного оборудования с использованием средств спутниковой навигации. Все ПЕ оснащены специальным оборудованием, которое дает, возможность в реальном времени наблюдать и более оперативно управлять движением городского маршрутизированного транспорта.

Автоматизированная навигационная система слежения за передвижением ПЕ предназначена для решения задач оперативного управления работой городского маршрутизированного транспорта, а также предоставления информационных услуг пассажирам посредством Интернет.

Основными целями навигационной системы являются:

- автоматизация работы центрального диспетчерского пункта ЕМУП ТТУ;
- создание средств информирования пассажиров о движении ПЕ.

Основными задачами навигационной системы являются:

- контроль текущего местоположения и текущего состояния ПЕ;
- контроль за соблюдением расписания движения ПЕ;
- оперативное управление движением ПЕ с помощью радиосвязи;
- помощь в оперативном разрешении штатных и внештатных ситуаций, возникающих в работе ПЕ;
- оперативный доступ пассажиров к информационно-справочной системе ЕМУП ТТУ в целях получения информации о прогнозируемом времени прибытия ПЕ на остановочный пункт;
- формирование отчетов.

Для решения задач автоматизированной навигационной системы слежения за передвижением ПЕ разработано автоматизированное рабочее место диспетчера. Внешний вид АРМа представлен на рис. 14.24.

Программа multifunctional и отвечает поставленным требованиям:

- отображение местоположения всех ПЕ на карте города в реальном времени;
- выбор и слежение за одним или несколькими ПЕ на карте города в реальном времени;
- группировка ПЕ по признаку:
 - ПЕ, находящиеся на заданном участке;
 - маршруты ПЕ, проходящие через заданный участок;
 - ПЕ заданного маршрута;
- фильтрация отображаемых ПЕ по признаку:
 - простаивающие;
 - находящиеся в движении;
 - на маршруте;
 - движущихся по обходному маршруту;

- принадлежащие к одной из групп (по виду транспорта, по депо, по маршруту);
- введение ограничения скорости на участках маршрутной сети;
- изменение масштаба отображаемого участка карты города;
- просмотр свойств ПЕ;
- просмотр информации о группе ПЕ;
- изменение статусов ПЕ;
- цветовое выделение заданного маршрута;
- протоколирование работы диспетчера;
- проверка прав доступа диспетчера по имени/паролю;
- хранение информации в единой базе данных;
- формирование следующих отчетов:
 - простои ПЕ заданного маршрута или группы;
 - корректность выполнения путевого листа заданной ПЕ;
 - итоги работы диспетчерской смены;
- формирование команд бортовым комплектам в соответствии с требованиями технологического процесса.

На рис. 14.25 показано отображение выбранного маршрута с выделением маршрута следования и выбранного ПЕ с привязкой к расписанию данного ПЕ.

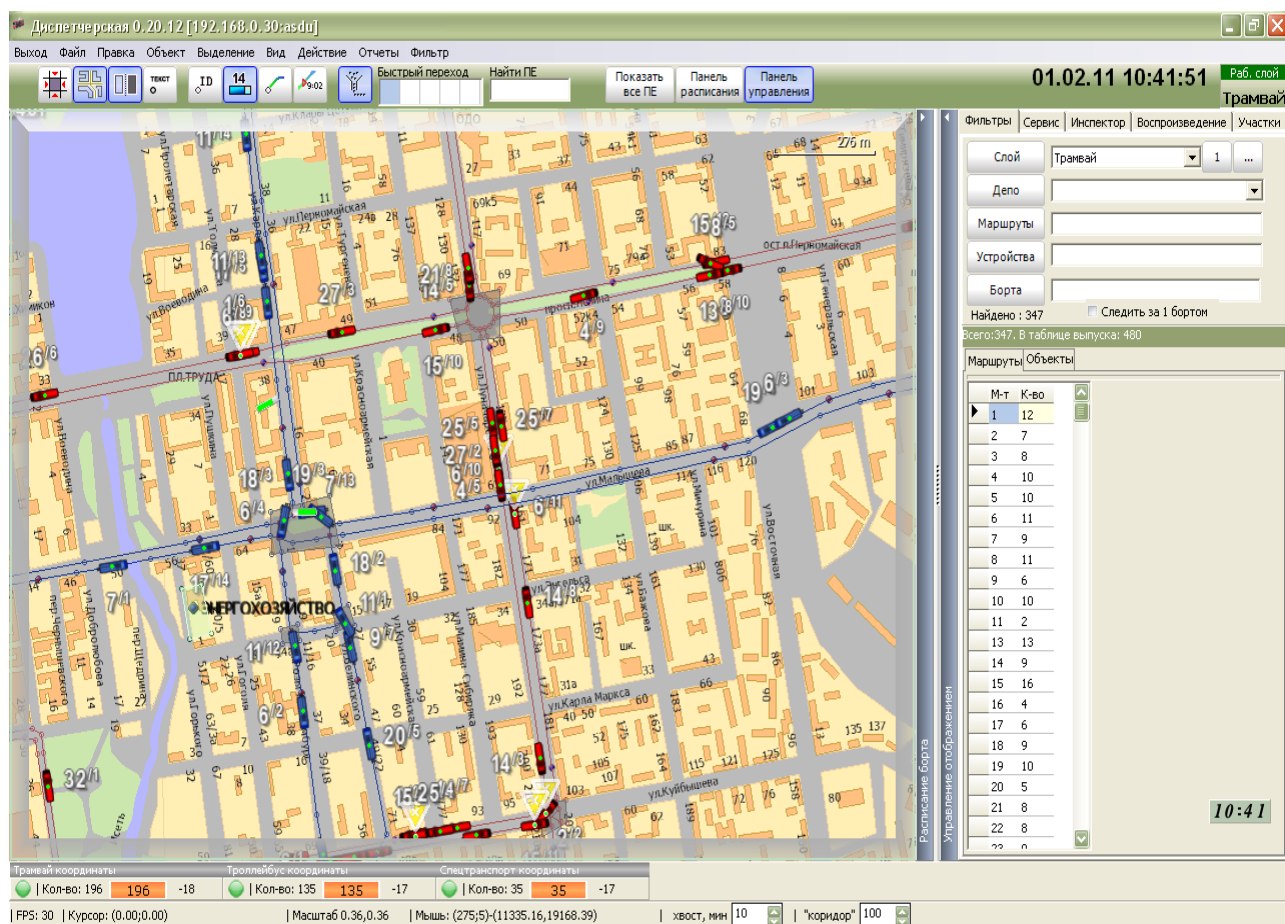


Рис. 14.24. Общий вид диспетчерской задачи

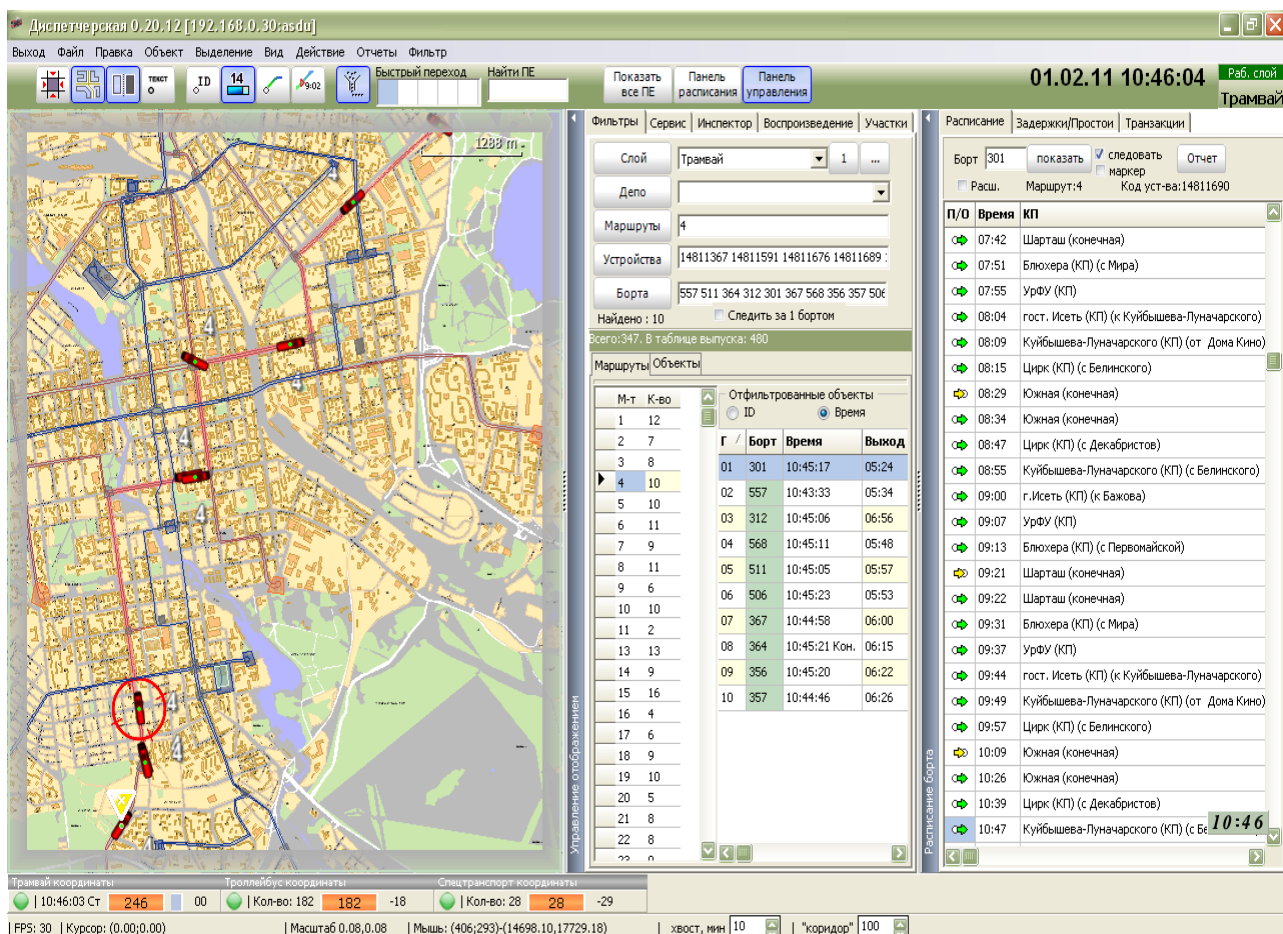


Рис. 14.25. Выбор маршрута

По выбранному ПЕ можно получить отчет о движении. Отчет содержит фактическое время прохождения ПЕ через остановки с привязкой к расписанию. Сразу отображаются рейсы и пробег ПЕ. При необходимости можно просмотреть трек движения ПЕ за выбранный период (вкладка «Воспроизведение»).

Таким образом, регулярность, ритмичность, бесперебойность движения транспортной сети как качественные показатели работы ЕМУП ТТУ отслеживаются диспетчером визуально в реальном времени, что позволяет оперативно управлять работой транспортной сети и оптимизировать процесс движения электротранспорта.

14.3. ДИСПЕРСИОННЫЙ АНАЛИЗ ТЕХНИКО-ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ ЛОГИСТИЧЕСКОЙ МОДЕЛИ ТРАНСПОРТНОЙ СЕТИ⁸

Воспользуемся однофакторным дисперсионным анализом (см. раздел 8.3) для проверки гипотез о наличии связи между откликом транспортной системы (доход предприятия) и исследуемыми факторами (рабочие дни, среднемесячная температура, простои транспорта), а также для установления силы влияния факторов и их взаимодействий. Анализ проведен укрупненно (по месяцам) и

⁸ Работа проводилась на кафедре автоматики и управления в технических системах ФГАОУ ВПО «УрФУ им. первого Президента России Б.Н. Ельцина» совместно со студентами Барковским Д., Шрайнер А.

детально (по дням). Выдвинута нулевая гипотеза о том, что фактор не влияет на оцениваемую характеристику. Расчетные значения дисперсионных соотношений сравниваются с табличным значением по Фишеру F_T , выбранного в зависимости от уровня значимости и степеней свободы. В случае превышения дисперсионного соотношения относительно табличного значения гипотеза отвергается, т. е. фактор влияет на оцениваемую характеристику.

Укрупненный анализ показал, что самым влиятельным фактором на доход является количество рабочих дней в месяце, в частности это проявилось на маршрутах троллейбусов №3, 5, 12, 17 и трамваев №1, 4, 5, 10, 14, 20, 21, 24, 25, 33. Анализ влияния среднемесячной температуры показал, что температура оказывает большее влияние на доход, в частности средняя температура сильно влияет на доход маршрутов троллейбусов №3, 5, 12, 17 и трамваев №11, 25. Укрупненный анализ показал, что фактор простоев электротранспорта влияет на доход маршрутов, проходящих по загруженным автомобильным транспортом улицам (троллейбус №7, трамваи №9, 23, 31).

Детальный анализ показал влияние календарных дней на доход на всех маршрутах трамвая (рис. 14.26). Определилась общая тенденция – ярко выраженная зависимость дохода от отпускных месяцев (июль, август, сентябрь и октябрь). Прослеживается влияние неучтенных факторов: введение социальных электронных карт в апреле, а также повышение оплаты проезда в октябре 2010 г. Кроме того, анализ показал существование маршрутов, на которых рабочие дни и отпускные месяцы слабо влияют на доход маршрута (трамваи № 10, 13). Это связано с наложением трудовых и культурно-бытовых потребностей населения на этих маршрутах.

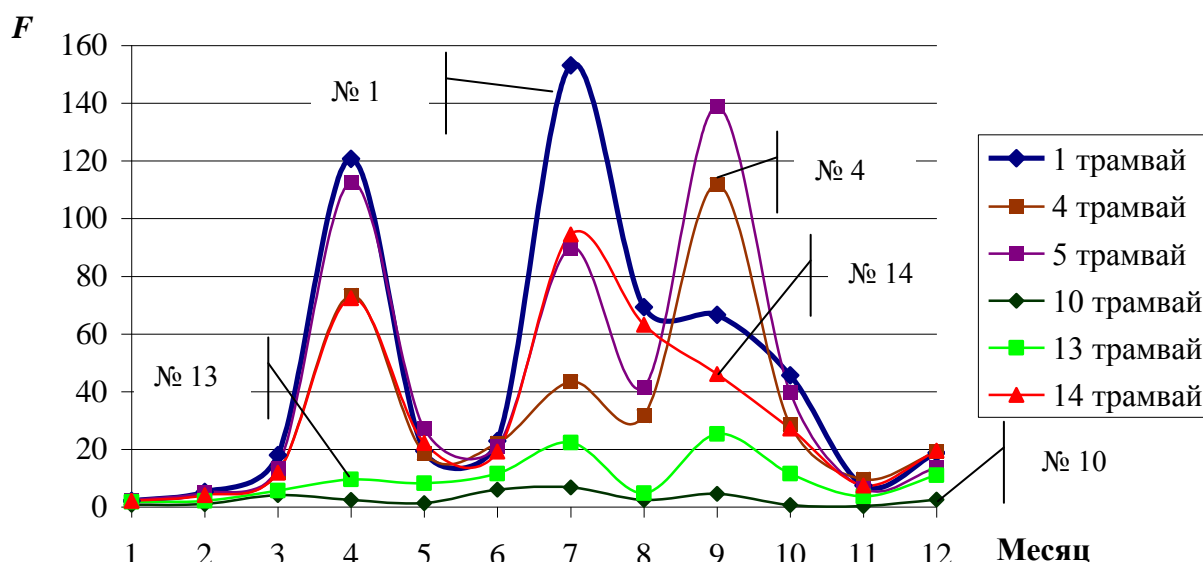


Рис. 14.26. Оценка влияния рабочих дней на доход

Детальный анализ влияния средней дневной температуры на доход показал слабую зависимость. Обнаружено влияние средней дневной температуры на доход для двух месяцев – мая и августа. Возможно, это связано с сезонными работами в садах и огородах и соответствующими выездами пассажиров на

окраины города (маршруты трамваев № 11, 24, 19, 10, 23, рис. 14.27). Кроме того прослеживается влияние неучтенных факторов: введение социальных электронных карт в апреле, особенно для маршрутов, проходящих через центр города (троллейбусы № 4, 13, 17, 3, трамваи № 26, 18, рис. 14.28).

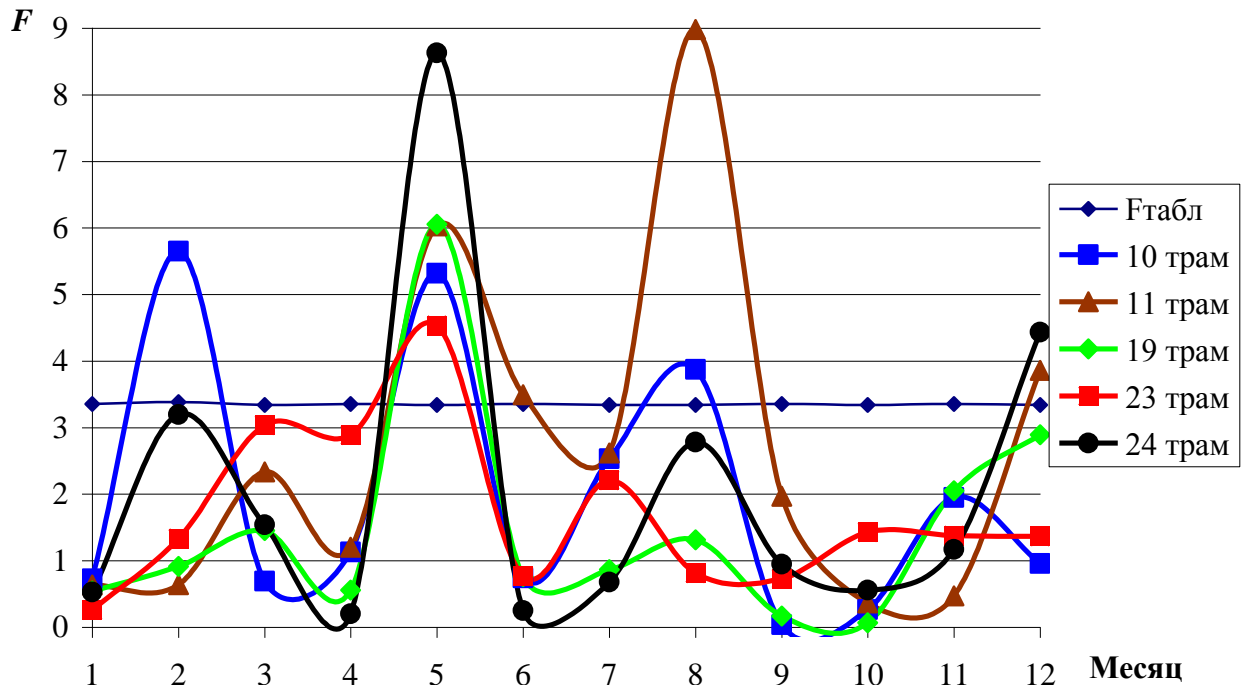


Рис. 14.27. Оценка влияния температуры на доход

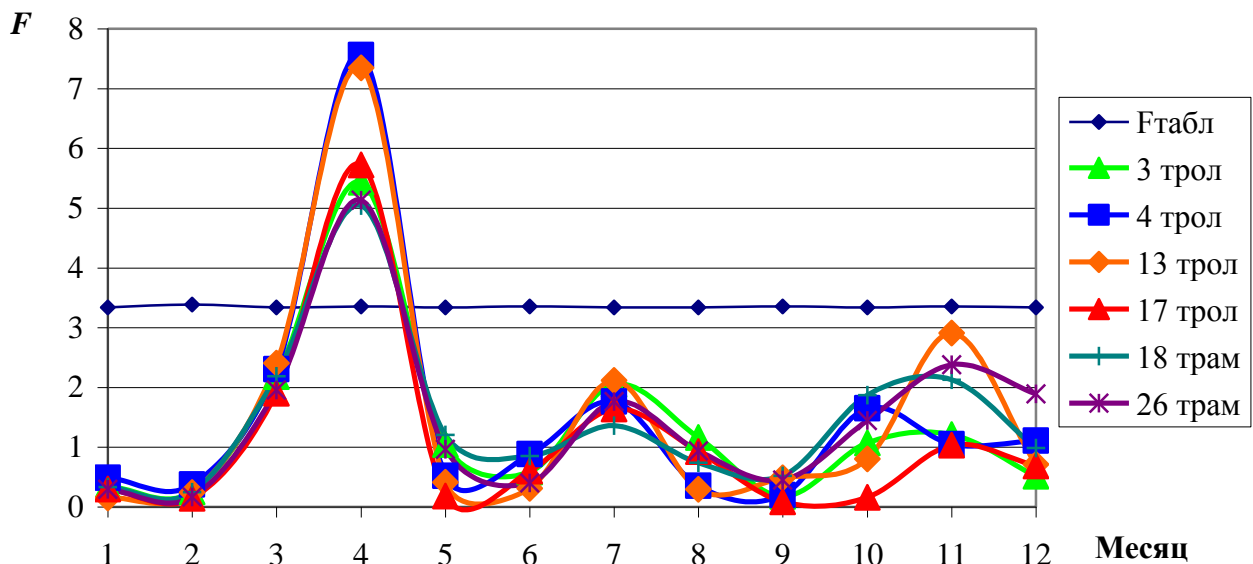


Рис. 14.28. Оценка влияния температуры на доход в апреле

Детальный анализ влияния простоев маршрутизированного городского электротранспорта на доход показал, что только в летние месяцы проявляется

слабая зависимость (рис. 14.29, 14.30). Наибольшая зависимость выявлена у троллейбусов №3, 12. Кроме того проявился парадокс – чем больше простаивает городской электротранспорт, тем больше доход по маршруту.

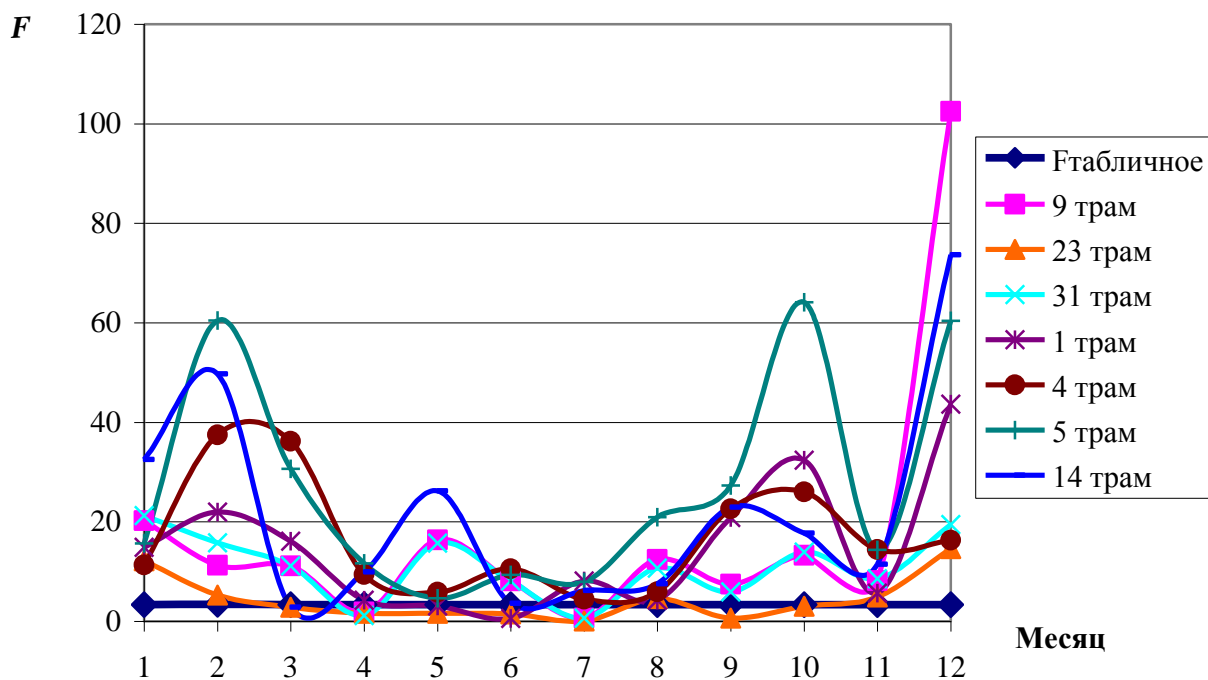


Рис. 14.29. Оценка влияния простоев транспорта на доход у трамваев

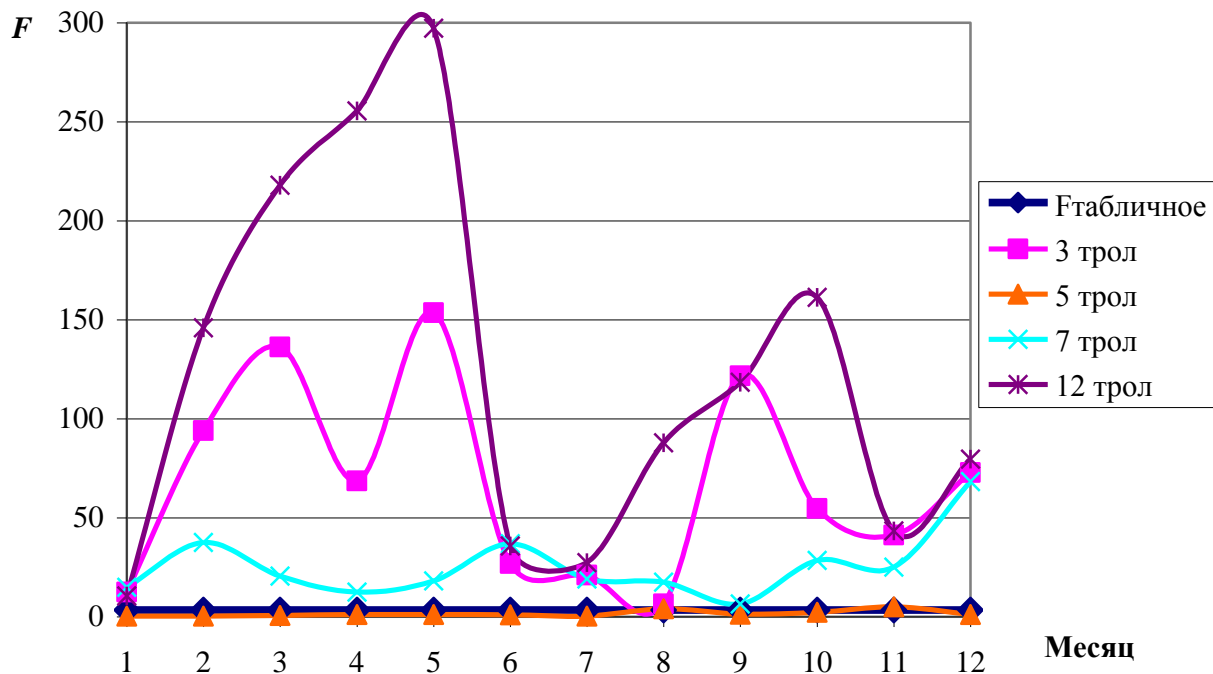


Рис. 14.30. Оценка влияния простоев транспорта на доход у троллейбусов

ЗАКЛЮЧЕНИЕ

Как отмечено во введении, учебное пособие посвящено математическому моделированию сложных вероятностных систем. Конечно, для решения задач статистического анализа имеется целый ряд учебников и учебных пособий, некоторые из них отмечены в библиографическом списке [6–8, 10–14, 36]. Целью настоящего учебного пособия являлось не только раскрыть ряд важных аспектов теории вероятностей и математической статистики в приложении к методам моделирования вероятностных систем, но и продемонстрировать на характерных (в авторских интерпретациях и построениях) моделях особенности математического моделирования и дать оценку важнейших характеристик рассмотренных моделей.

Естественно, следующим шагом в раскрытии возможностей вероятностных моделей является их использование в задачах прогнозирования и оптимального управления (в рамках структурной или параметрической оптимизации). При этом становится необходимым использовать комплексный критерий, по оценке которого можно связать, например, разноразмерные параметры системы, а также обеспечить заданную вероятность выполнения технических требований для этих параметров при действии различных случайных возмущений [37]. Алгоритмы оптимизации (поиск максимума или минимума поверхности реакции в некотором пространстве факторов) могут быть различны, а используемые вероятностно-статистические методы, как правило, опираются на теорию вероятности и математическую статистику с использованием различных критериев и доверительных интервалов.

Данное учебное пособие в определенной мере является подготовительным этапом в решении более сложных задач управления.

Рассмотренный в данном учебном пособии круг вопросов касается вероятностных подходов к проблеме математического моделирования в сложных системах. Естественно, это не умаляет роль детерминированных подходов к разработке и анализу математических моделей, в развитии которых авторам также принадлежит определенная роль. Комбинированные вероятностно-детерминированные модели с применением методов идентификации занимают достойное место в решении актуальных задач оптимизации и управления. Эти проблемы освещены в ряде публикаций [35, 38–42].

ОТВЕТЫ НА КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Модель – это физический или абстрактный образ моделируемого объекта, удобный для проведения исследований и позволяющий адекватно отображать интересующие исследователя физические свойства и характеристики объекта.

2. Понятие «цель» трансформируется от идеальных устремлений к материальному воплощению, конечному результату деятельности.

3. Р. Шеннон дает следующее определение: «имитационное моделирование есть процесс конструирования реальной системы и постановки экспериментов на этой модели с целью оценить (в рамках ограничений, накладываемых некоторым критерием или совокупностью критериев) различные стратегии, обеспечивающие функционирование данной системы».

4. К первому относятся тесты, в которых задаются правдоподобные значения входов и управляющих воздействий. Если в этом случае расчеты по модели не противоречат известным законам поведения реального объекта, то это говорит в пользу адекватности модели. В противном случае необходимо найти причины несогласованности и перестроить модель. Второй тип тестов основан на использовании критических ситуаций, т. е. данных, которые не характерны для исследуемого объекта, но тем не менее могут иметь место. Эти тесты особенно важны для сложных моделей, предназначенных для долгосрочного прогнозирования, поскольку чем лучше модель описывает поведение объекта в критических условиях, тем больше можно быть уверенным в правильности расчетов для нормальных условий.

5. В теоретико-вероятностной математической модели можно точно указать законы распределения случайных величин, а в статистической заранее нельзя указать законы распределения случайных величин.

6. $[M(X)] = [X]$, $[D(X)] = [X]^2$, $[\sigma(X)] = [X]$, $[A(X)] = 1$, $[E(X)] = [X]$, $\text{cov}(X, Y) = [X] \cdot [Y]$, $r(X, Y) = 1$.

7. Если функция неотрицательная с конечной площадью под графиком, то после нормирования она может являться плотностью распределения. Если функция не убывает, ограничена сверху и в пределе $\lim_{x \rightarrow -\infty} f(x) = 0$, после нормирования она может быть функцией распределения.

8. Если X не зависит от Y , то и Y не зависит от X . Если X и Y независимы, то они некоррелированы. Другими словами, если X и Y коррелированы, то они зависимы. Обратное утверждение неверно, т.е. X и Y могут быть зависимыми и некоррелированными.

9. Если $r(x, y) > 0$, то линия тренда возрастает. Если $r(x, y) < 0$, то линия тренда убывает. Если $r(x, y) = 0$, то либо линия тренда горизонтальна, либо она отсутствует, и в этом случае требуются дополнительные исследования.

10. Ковариация, деленная на произведение стандартных отклонений, есть безразмерный коэффициент связи – коэффициент корреляции. Иначе ковариация нормированных отклонений есть коэффициент корреляции.

11. Площадь под графиком для функции плотности распределения хи-квадрат равна 1.

12. Закон редких явлений называется законом Пуассона, который выражает биномиальное распределение при большом числе опытов и малой вероятности события.

13. Распределение Эрланга k -го порядка является универсальным распределением. Путем изменения целочисленного параметра k можно трансформировать данное распределение случайной величины в одно из известных. Например, при $k = 1$ это распределение трансформируется в экспоненциальное, при $k = 2$ – в распределение Релея, при $k > 10$ – в нормальное распределение. Выбор того или иного распределения, а следовательно, задание параметра k , определяется результатами обработки статистических данных.

14. Функция выборочной плотности является гистограммой частот. Поэтому сумма всех значений этой функции на ее области определения должна равняться 1. В то же время сумма значений гистограммы частот равняется количеству элементов выборки.

15. Однопараметрическим семейством является показательное распределение, двухпараметрическим – нормальные и равномерно распределенные величины. Треугольные распределения имеют три параметра распределения. Трапециевидные распределения имеют четыре параметра. При сложении двух параметрических семейств количество параметров, как правило, увеличивается.

16. Функция распределения $F(x)$ изменяется от 0 до 1 и убывает. Функция плотности распределения $f(x)$ неотрицательна, и площадь ее под графиком равна 1.

17. По функции $f(x, y)$ построим плотность предельного распределения $f(x)$. Генерируем случайную величину \tilde{x} по закону $f(x)$. Строим плотность условного распределения $f(y \mid X = \tilde{x})$. Генерируем случайную величину \tilde{y} по закону $f(y \mid X = \tilde{x})$. В итоге получаем (\tilde{x}, \tilde{y}) .

18. В ходе эксперимента изменяли дисперсию какого-либо параметра, а в качестве результатов модели анализировали математическое ожидание выходной характеристики. Если математическое ожидание выходной характеристики редко попадает в критическую область, то можно сделать вывод, что она почти нечувствительна к изменению дисперсии входного параметра.

19. В ходе эксперимента изменяли математическое ожидание какого-либо параметра, а в качестве результатов модели анализировали дисперсию выходной характеристики. Если дисперсия выходной характеристики часто попадает в критическую область, то можно сделать вывод, что она очень чувствительна к изменению математического ожидания входного параметра.

20. Если коэффициент корреляции равен +1, то все точки двумерной выборки лежат на возрастающей прямой. Если коэффициент корреляции равен -1, то все точки двумерной выборки лежат на убывающей прямой.

21. Перед проведением дисперсионного анализа рекомендуется проверить следующие гипотезы. Групповые результативные признаки должны:

- 1) быть независимы, т.е. межгрупповые корреляции должны равняться нулю (см. раздел 7.3);
- 2) быть нормально распределены в пределах групп (критерий Пирсона);
- 3) иметь одинаковые дисперсии (критерий Бартлетта).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Штофф В.А. Моделирование и философия / В.А. Штофф. М.: Наука, 1966. 301 с.
2. Клаус Г. Кибернетика и философия / Г. Клаус. М.: Изд-во иностр. лит-ры, 1963. 262 с.
3. Уёмов А.И. Логические основы метода моделирования / А.И. Уёмов. М.: Мысль, 1971. 311 с.
4. Новый энциклопедический словарь. М.: Большая Российская энциклопедия : РИПОЛ КЛАССИК, 2004. 1456 с.
5. Новик И.Б. Философские вопросы моделирования психики / И.Б. Новик. М.: Наука, 1969. 174 с.
6. Тарасик В.П. Математическое моделирование технических систем : учебник для вузов / В.П. Тарасик. М.: Наука, 1997. 600 с.
7. Волкова В.Н. Основы теории систем и системного анализа : учебник для студентов вузов, обучающихся по специальности «Системный анализ и управление» / В.Н. Волкова, А.А. Денисов. СПб.: СПбГТУ, 1997. 510 с.
8. Альянах И.Н. Моделирование вычислительных систем / И.Н. Альянах. Л.: Машиностроение, 1988. 233 с.
9. Канторович Л.В. Математические методы организации и планирования производства / Л.В. Канторович. Л.: ЛГУ, 1939. 69 с.
10. Владимировский Б.М. Математика. Общий курс : учебник для вузов / Б.М. Владимировский, А.Б. Горстко, Я.М. Ерусалимский. СПб.: Лань, 2002. 960 с.
11. Шеннон Р. Имитационное моделирование систем – искусство и наука / Р. Шеннон. М.: Мир, 1978. 308 с.
12. Компьютерный анализ и интерпретация эмпирических зависимостей : учебник / С.В. Поршневу [и др.]. М.: Бином-Пресс, 2009. 336 с.
13. Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников / А.И. Кобзарь. М.: ФИЗМАТЛИТ, 2006. 816 с.
14. Письменный Д.Т. Конспект лекций по теории вероятностей и математической статистике / Д.Т. Письменный. М.: Айрис-пресс, 2004. 256 с.
15. Mandelbrot B.B. Les Objets Fractals: Forme, Hasard et Dimension / B.B. Mandelbrot. Paris: Flammarion, 1975. 412 p.
16. Helge van Koch. Sur une courbe continue sans tangente, obtenue par une construction geometrique elementaire / Van Koch Helge // Arkiv for Matematik. 1904. № 1. P. 681 – 704.
17. Хомский Н. Формальные свойства грамматики / Н. Хомский // Кибернетический сборник. М., 1966. Вып. 2. С. 121 – 130.
18. Prusinkiewicz P. The Algorithmic Beauty of Plants / P. Prusinkiewicz, A. Lindenmeyer. New-York: Springer-Verlag, 1996. 228 p.
19. Трофимов С.П. Обобщенные многоугольники и обработка графической информации, представленной в форме несвязанных ломаных линий / С.П. Трофимов, О.Е. Лейзеров // Информационные системы и технологии в

- радиотехнике, связи, автоматике и управлении. Вестник УГТУ-УПИ. Серия радиотехническая. Екатеринбург: УГТУ-УПИ, 2005. № 17(69). С. 85 – 90.
20. Водопьянов И.О. Моделирование и распознавание узоров на плоскости / И.О. Водопьянов, С.П. Трофимов, И.А. Буцырин // Науч. тр. междунар. науч.-практической конференции «СВЯЗЬ-ПРОМ 2008» в рамках V Евро-Азиатского международного форума «СВЯЗЬ-ПРОМЭКСПО 2008». Екатеринбург: Компания Реал-Медиа, 2008. С. 302 – 304.
21. Водопьянов И.О. Моделирование и распознавание регулярных линейчатых геометрических узоров на плоскости / И.О. Водопьянов, С.П. Трофимов // Науч. тр. 8-й Рос. конф. с междунар. участием «Новые информационные технологии в исследовании сложных структур». ISAM'10. Томск: ТГУ, 2010. С. 210.
22. Гренандер У. Лекции по теории образов: Анализ образов / пер. с англ.: У. Гренандер. М.: Мир, 1981. 448 с.
23. Decagonal and Quasi-Crystalline Tilings in Medieval Islamic Architecture / Peter J. Lu et al. // Science. 2007. № 315. P. 1106 – 1110.
24. Абдуллаева Г.Г. Распознавание и идентификация узоров азербайджанских ковров ручной работы / Г.Г. Абдуллаева, Н.Г. Курбанова, А.К. Кязим-заде // Изв. НАНА, серия «Информатика и проблемы управления». Баку, 2007. Т. XXVII, № 2-3. С. 23 – 28.
25. Graphene at the Edge: Stability and Dynamics / Çaglar Ö. Girit et al. // Science. 2009. № 323. P. 1705 – 1708.
26. Мотанов В.Г. Решётки на множестве $Z \times Z$, инвариантные относительно аффинных поворотов, образующих циклическую группу 6-го порядка / В.Г. Мотанов // Сб. науч. тр. М.: МГУП, 2005. С. 55 – 60.
27. Малеев А.В. Математическое моделирование кристаллических и квазикристаллических структур: автореф. дис. ... д-ра физ.-мат. наук / А.В. Малеев. Н. Новгород: НГУ, 2011. 50 с.
28. Сборник олимпиадных задач для специальности «Вычислительные машины, комплексы, системы и сети» : учеб. пособие / под ред. С.В. Поршнева. М.: КНОРУС, 2010. 280 с.
29. Дружинина Н.Г. Программный комплекс по составлению расписания маршрутизированного транспорта / Н.Г. Дружинина. М.: ВНИИЦ, 2005. № 50200501360.
30. Дружинина Н.Г. Программный комплекс по обработке путевого листа автотранспортной службы / Н.Г. Дружинина. М.: ВНИИЦ, 2006. № 50200600033.
31. Дружинина Н.Г. Диспетчер выпуска и движение подвижной единицы / Н.Г. Дружинина. М.: ВНИИЦ, 2006. № 5020061544.
32. Дружинина Н.Г. Транспорт города Екатеринбурга / Н.Г. Дружинина, О.Г. Трофимова. М.: ВНИИЦ, 2008. № 50200800285.
33. Табель учета рабочего времени водителей и кондукторов : свидетельство о регистрации электронного ресурса. № 15296. 05.02.2010 / Н.Г. Дружинина, О.Г. Трофимова. М.: ВНИИЦ, 2010. № 50201000268.

34. Моделирование систем с использованием информационных технологий : учеб. пособие / В.Г. Лисиенко, Н.Г. Дружинина, О.Г. Трофимова, С.П. Трофимов. Екатеринбург: УГТУ-УПИ, 2009. 440 с.
35. Дружинина Н.Г. Информационно-коммуникационная транспортная система / Н.Г. Дружинина, О.Г. Трофимова // Дистанционное и виртуальное обучение. 2009. № 9 (27). С. 77 – 94.
36. Орлов А.И. Теория принятия решений : учеб. пособие / А.И. Орлов. М.: Март, 2004. 656 с.
37. Шатунов П.И. Многопараметрическая оптимизация сложных нелинейных нестационарных систем управления подвижными объектами на примере сложной системы управления расходом топлива / П.И. Шатунов, Г.П. Лосев // Новые методы передачи и обработки информации : сб. результатов науч. исслед. сотрудников радиотехнического фак. УГТУ-УПИ. Екатеринбург: УГТУ-УПИ, 2005. С. 248 – 257.
38. Лисиенко В.Г. Альтернативная металлургия: проблема легирования, модельные оценки эффективности / В.Г. Лисиенко, Н.В. Соловьева, О.Г. Трофимова; под ред. В.Г. Лисиенко. М.: Теплотехник, 2007. 440 с.
39. Лисиенко В.Г. Моделирование объектов с распределенными параметрами на примере трехуровневых АСУ нагревом материала : учеб. пособие / В.Г. Лисиенко, З.Г. Салихов, О.А. Гусев. Екатеринбург: УГТУ-УПИ; М.: МИСиС, 2004. 163 с.
40. Моделирование и автоматизированное управление / В.Г. Лисиенко [и др.]; под ред. В.Г. Лисиенко. Новоуральск: НИЯУ-МИФИ-НТИ, 2010. 179 с.
41. Зобнин Б.Б. Моделирование систем / Б.Б. Зобнин. Екатеринбург: УГГГА, 2001. 128 с.
42. Самарский А.А. Математическое моделирование. Идеи. Методы. Примеры / А.А. Самарский, А.П. Михайлов. М.: Наука ФИЗМАТЛИТ, 1997. 320 с.

Учебное издание

Лисиенко Владимир Георгиевич

Трофимова Ольга Геннадиевна

Трофимов Сергей Павлович

Дружинина Надежда Геннадьевна

Дюгай Павел Алексеевич

МОДЕЛИРОВАНИЕ СЛОЖНЫХ ВЕРОЯТНОСТНЫХ СИСТЕМ

Редактор *И.В. Кошунова*

Компьютерный набор *О.Г. Трофимовой, Н.Г. Дружининой*

Подписано в печать 09.12.2011. Формат 60х84 1/16.

Бумага писчая. Плоская печать. Усл. печ. л. 11,63.

Уч.-изд. л. 14,3. Тираж 50 экз. Заказ

Редакционно-издательский отдел УрФУ

620002, Екатеринбург, ул. Мира, 19

rio@mail.ustu.ru

Ризография НИЧ УрФУ

620002, Екатеринбург, ул. Мира, 19